Display II

CANADIAN JOURNAL OF MATHEMATICS

Journal Canadien de Mathématiques

VOL. VII - NO. 1
1955

Published for

THE CANADIAN MATHEMATICAL CONGRESS

by the

University of Toronto Press

The chief languages of the *Journal* are English and French.

Manuscripts for publication in the *Journal* should be sent to the *Editor-in-Chief*, H. S. M. Coxeter, University of Toronto. Everything possible should be done to lighten the task of the reader; the notation and reference system should be carefully thought out. Every paper should contain an introduction summarizing the results as far as possible in such a way as to be understood by the non-expert.

All other correspondence should be addressed to the *Managing Editor*, G. de B. Robinson, University of Toronto.

The *Journal* is published quarterly. Subscriptions should be sent to the *Managing Editor*. The price per volume of four numbers is $8.00. This is reduced to $4.00 for individual members of recognized Mathematical Societies.

The Canadian Mathematical Congress gratefully acknowledges the assistance of the following towards the cost of publishing this Journal:

University of British Columbia

Carleton College
Université Laval
University of Manitoba
McMaster University
Queen's University
St. Mary's University
Ecole Polytechnique
Loyola College
McGill University
Université de Montréal
Royal Military College
University of Toronto

National Research Council of Canada
and the
American Mathematical Society

AUTHORIZED AS SECOND CLASS MAIL, POST OFFICE DEPARTMENT, OTTAWA

COMBINATORIAL RELATIONS AND CHROMATIC GRAPHS

R. E. GREENWOOD AND A. M. GLEASON

1. Introduction. The following elementary logical problem was a question in the William Lowell Putnam Mathematical Competition held in March 1953 **(1)**:

> Six points are in general position in space (no three in a line, no four in a plane). The fifteen line segments joining them in pairs are drawn, and then painted, some segments red, some blue. Prove that some triangle has all its sides the same color.

This problem may be generalized as follows:

> N vertex points are given, and r colors are available. Each of the $\frac{1}{2}N(N-1)$ segments joining two vertices is colored by one and one only of the r given colors. Find the minimum integer $n = n(k_1, k_2, \ldots, k_r)$ such that, if $N > n$, either there will exist a subset of k_1 vertices with all interconnecting segments of the first color, or there will exist a subset of k_2 vertices with all interconnecting segments of the second color, ..., or there will exist a subset of k_r vertices with all interconnecting segments of the rth color.

The elementary problem is then the special case $r = 2$, $k_1 = 3$, $k_2 = 3$, and essentially asserts that $n(3, 3) < 6$. A slightly more general form would assert that $n(3, 3) = 6$.

THEOREM 1. $$n(3, 3) = 6.$$

Proof. To show $n(3, 3) > 5$, consider the five points to be the vertices of a regular pentagon. Let the two colors be red and blue. Color the edges red and the interior diagonals blue. In this assemblage of 5 vertices and 10 segments there is no triangle with all its sides the same color.

To show 6 is a sufficient number of vertices, select some one vertex point and call it a. Of the five line segments terminating at a, select a set of three all of which have the same color. Consider the three segments joining their farther ends in pairs. If no one of these segments is of the same color as the initial set, then all three segments are of the other color and do form a required triangle.

Note. The second part of the above proof is essentially the same as that given by one of the Putnam examination candidates.

2. General case for 2 colors. We define

$$n(1, m) = 1 \qquad (m = 1, 2, 3, \ldots),$$
$$n(k, 1) = 1 \qquad (k = 1, 2, 3, \ldots).$$

Received July 25, 1953; in revised form June 8, 1954.

It is easy to see that

$$n(2, m) = m \qquad (m = 2, 3, 4, \ldots),$$
$$n(k,\ 2) = k \qquad (k = 2, 3, 4, \ldots).$$

The relation $n(k, m) = n(m, k)$ is an obvious consequence of the symmetry or duality of the problem.

It is possible to establish two inequalities for $n(k, m)$ which are similar in appearance to a well-known identity for binomial coefficients.

For convenience, a set of vertices K will be said to be a red interconnected set if all segments joining pairs of vertices in K are colored red. A set of vertices M will be said to be a blue interconnected set if all segments joining pairs of vertices in M are colored blue. In the minimum vertex number $n(k, m)$, the first argument will be used for the red color, and the second argument for the blue color.

THEOREM 2. $$n(k, m) \leqslant n(k-1, m) + n(k, m-1).$$

Proof. Let T be any set of $n(k-1, m) + n(k, m-1)$ vertices. Form a chromatic graph by coloring each interconnecting segment either red or blue. Select one vertex and call it a. For vertex a consider the two associated numbers:

n_1 = number of vertices such that the segments joining them to a are red,
n_2 = number of vertices such that the segments joining them to a are blue.

Call these sets of vertices T_1 and T_2.

Since all vertices other than a belong to either T_1 or T_2, one obtains

$$n_1 + n_2 + 1 = n(k-1, m) + n(k, m-1).$$

If $n_1 < n(k-1, m)$, then $n_2 \geqslant n(k, m-1)$ and one selects the set T_2. In T_2 there is either a red interconnected set S_1 with k vertices or a blue interconnected set S_2 with $m - 1$ vertices. If the latter holds, the set $\{S_2 + \text{vertex } a\}$ is a blue interconnected set with m vertices. Hence the inequality of Theorem 2 is established for this case.

If $n_1 \geqslant n(k-1, m)$, one selects the set T_1. In T_1 there is either a red interconnected set S_3 with $k - 1$ vertices or a blue interconnected set S_4 with m vertices. In the first case the set $\{S_3 + \text{vertex } a\}$ is seen to be a red interconnected set with k vertices. The proof of Theorem 2 is now complete.

COROLLARY 1.

$$n(k, m) \leqslant \binom{k+m-2}{k-1}.$$

By an obvious extension of this argument one may prove:

COROLLARY 2. *For the case of more than two colors,*

$$n(k_1, k_2, \ldots, k_r) \leqslant n(k_1 - 1, k_2, \ldots, k_r) + n(k_1, k_2 - 1, \ldots, k_r) + n(k_1, k_2, \ldots, k_r - 1).$$

COROLLARY 3. *For the multi-color case, an inequality on the size of the minimum number is afforded by the multinomial coefficient*

$$n(k_1 + 1, k_2 + 1, \ldots, k_r + 1) \leqq \frac{(k_1 + k_2 + \ldots + k_r)!}{k_1!\, k_2! \ldots k_r!}.$$

THEOREM 3. *If* $n(k - 1, m) = 2p$ *and* $n(k, m - 1) = 2q$, *then*

$$n(k, m) < 2p + 2q = n(k - 1, m) + n(k, m - 1).$$

Proof. Take a set of $2p + 2q - 1$ vertices. Select one vertex and call it a. There are $2p + 2q - 2$ segments terminating at a. Three possibilities might arise:

(a) $2p$ or more segments terminating at a are red,

(b) $2q$ or more segments are blue, or

(c) $2p - 1$ segments are red and $2q - 1$ segments are blue.

For case (a) consider the set T_1 of the vertices at the farther ends of the $2p$ or more red segments. Since the number of vertices in T_1 is greater than or equal to $n(k - 1, m)$, there is either a red interconnected set S_1 with $k - 1$ vertices or a blue interconnected set S_2 with m vertices. In the first situation, the set $\{S_1 + \text{vertex } a\}$ is a red interconnected k vertex set. Thus the theorem is true for case (a).

Likewise, a similar argument holds for case (b).

Case (c) cannot hold for each vertex of the chromatic graph. For if it did hold, then there would be $(2p + 2q - 1)(2p - 1)$ red ends. This calls for an odd number of red ends, but since each segment has two ends, the number of red ends is required to be even. Hence there must be at least one vertex where either case (a) or case (b) holds, and the theorem is true for both these cases.

3. Special values for the two color problem. It is known that $n(2, 4) = 4$ and $n(3, 3) = 6$. From Theorem 3 it follows that $n(3, 4) \leqq 9$. This value shows that the strict inequality sign in Corollary 1 has to be used in some cases.

To assist in evaluating $n(3, 5)$ consider the 13 element field with the elements numbered from 0 to 12 inclusive and take each field element as a vertex in a graph. The cubic residues in this field are 1, 5, 8, and $12 \equiv -1 \pmod{13}$. If the difference of two vertices is a cubic residue, color the corresponding line segment red. If the difference is not a cubic residue, color the corresponding line segment blue. Since 1 and -1 are both cubic residues, the order of differencing is immaterial. In this chromatic graph there is no subset of three which is red interconnected and no subset of five which is blue interconnected. Hence $n(3, 5) > 13$.

Theorem 2 may now be used to remove the doubt as to the values $n(3, 4) \leqq 9$ and $n(3, 5) \geqq 14$. Since $n(2, 5) = 5$ and $n(3, 5) \leqq n(2, 5) + n(3, 4)$, one readily obtains that $n(3, 4) = 9$ and $n(3, 5) = 14$.

It will now be shown that $n(4, 4) > 17$. Consider the field of 17 elements numbered from 0 to 16 inclusive. Let each element be a vertex. If two vertices

have a numerical difference which is a quadratic residue of 17, the corresponding segment is to be colored red. Note that $-1 \equiv 16 \pmod{17}$, a square, and hence the order in which the subtraction is performed makes no difference. The other segments are to be colored blue.

Suppose that some four vertices are connected by segments of one color. Without loss of generality, one of these vertices may be considered to be the field element marked 0. Call the other three a, b, c. Then the six numbers a, b, c, $a-b$, $a-c$, and $b-c$ are either all residues or all non-residues. Since a is not the zero element in the field, multiplication by a^{-1} is permissible. If one sets $B = ba^{-1}$ and $C = ca^{-1}$, one can consider the new set of six numbers $1, B, C, 1-B, 1-C, B-C$. All of these must be quadratic residues of 17, further none of them can be zero. But the residues of 17 are 1, 2, 4, 8, 9, 13, 15 and 16 and it is impossible for all members of the set above to be residues. Theorem 3 may now be used to establish that $n(4, 4) = 18$.

On the basis of these values, an array may be set up giving the known values $n(k, m)$ as entries in the body of Table 1.

TABLE I

$m =$	1	2	3	4	5	6	7
$k = 1$	1	1	1	1	1	1	1
2	1	2	3	4	5	6	7
3	1	3	6	9	14	19 or less	
4	1	4	9	18	31 or less		
5	1	5	14	31 or less			
6	1	6	19 or less				
7	1	7					

The authors desire to point out that deep combinatorial questions are encountered in attempting evaluations of the minimum numbers. In this paper proof or disproof of the possibility of a chromatic graph construction leans heavily upon finite field theory.

4. Three or more colors. For three colors, red, green and blue, the evaluation of $n(3, 3, 3)$ may be of interest.

THEOREM 4. $n(3, 3, 3) = 17$.

Proof. It will be convenient to give the proof in two parts. First, one may show that $n(3, 3, 3) > 16$ by use of a $2^4 = 16$ element field. In such a field

the elements may be identified by the symbols 0, x, $x^2, \ldots, x^{15} = 1$. To assist in identifying cubic residues, one may factor

$$x^{15} - 1 \equiv (x^4 - x - 1)(x^{11} + x^8 + x^7 + x^5 + x^3 + x^2 + x + 1) \pmod{2},$$

and take x as a root of $x^4 - x - 1 \equiv 0 \pmod 2$. Then the cubic residues of the field elements are x^3, $x^3 + x^2$, $x^2 + x$, $x^3 + x^2 + x + 1$, and 1. The field elements are taken as vertices and a chromatic graph is constructed as follows. If the difference of two vertex elements is a cubic residue, the corresponding segment is colored red. If the difference of two elements belongs to the first coset of the cubic residues in the multiplicative group of the non-zero field elements, the segment is colored green. If the difference belongs to the second coset, the segment is colored blue. These specifications define completely the chromatic graph. Again one notes that $-1 \equiv 1 \pmod 2$, and the order of differencing is immaterial.

Suppose now that a set of three vertices is completely interconnected by segments of the same color. Without any loss of generality, the color may be considered as red and the vertices may be considered as having been numbered 0, 1 and A, where A and $(1 - A)$ are both cubic residues. But a comparison with the list of cubic residues above shows that this situation cannot arise. Hence there is no set of three vertices interconnected by segments of the same color, and $n(3, 3, 3) > 16$.

For the second part, it must be proved that 17 is a sufficient number. In any chromatic graph with 3 colors and 17 vertices, select any one vertex and call this vertex a. At least six segments terminating at a must all have the same color. If in the subgraph consisting of a and the six farther ends there is no triangle of the original color, then all interconnections among the six farther ends must be of the other two colors. But this is now a case of $n(3, 3) = 6$, and there is an interconnected set of three here.

Remark. This example serves also to prove that $n(6, 3) > 16$. One imagines the graph to be examined by a man who is red-green color blind and hence sees only two colors, namely blue and red-green. Supposing that there is no set of three vertices interconnected in blue, then the existence of a set of six interconnected in red-green would furnish a red triangle or a green triangle, since $n(3, 3) = 6$.

THEOREM 5. $41 < n(3, 3, 3, 3) < 66.$

Proof. The upper bound given in the theorem follows at once in the same way as used in the proof of Theorem 4, i.e., reducing to the next lower case.

To establish a lower bound, one could consider chromatic graphs whose cosets of fourth power residues were employed in connection with fields in which the multiplicative group is of order $4k$. This suggests the primes 41, 53, 61 and the extension field of 49 elements. Unfortunately, -1 is not a quartic residue for the cases 53 and 61, and the order of subtraction of vertex

elements would affect the question of what coset of residues the difference is in, and thus lead to a non-unique coloring scheme. For the 41 element field, the quartic residues are 1, 4, 10, 16, 18, 23, 25, 31, 37 and 40. There are no consecutives in this list, and thus a triangle whose vertices are 0, 1, and A cannot be interconnected by segments of the same color. Thus $n(3, 3, 3, 3) > 41$. The authors have not further narrowed the range on $n(3, 3, 3, 3)$.

5. Upper and Lower Bounds for $n(3^r)$. Let $t_r = n(3^r) = n(3, 3, \ldots, 3)$. The upper bounds used in Theorems 1, 4 and 5 can all be obtained by the use of

THEOREM 6. $\quad t_{r+1} \leqq (r+1)(t_r - 1) + 2.$

This theorem is easily proved by induction; and then it is trivial to establish, also by induction, that

$$t_{r+1} \leqq 3(r+1)!.$$

A somewhat sharper inequality may be obtained, however, without any added difficulties. It has already been established that

$$t_r \leqq [(r!)\, e] + 1, \qquad r = 2, 3, 4,$$

where $[M]$ means the greatest integer contained in M. Such a bound holds for all integers $r \geqq 2$, for if it did not there would be a least integer, say $s + 1$, for which the relation failed to hold. By Theorem 6,

$$t_{s+1} \leqq (s+1)[(s!)\, e] + 2, \qquad s \geqq 2.$$

But $[(s+1)!\, e] = (s+1)[(s!)\, e] + 1$, and hence

$$t_{s+1} \leqq [(s+1)!\, e] + 1$$

and the stated upper bound follows.

For t_5, an upper bound of 327 is thus obtained. In order to determine a possible lower bound, one may generalize to some extent the previous arguments for lower bounds.

Assume a chromatic graph with p vertices (where the p vertices are to be thought of as field elements) and 5 colors. To ensure uniqueness of coloring, one would require that -1 be a 5th power residue. If p is odd, this means that 5 divides $\frac{1}{2}(p-1)$. If p is even, then $-1 \equiv 1 \pmod 2$ and the condition that -1 be a 5th power residue is trivially satisfied.

For $r = 5$, and restricting p to primes less than 327, one sees that 317, 313, 307, 293, 283 do not satisfy the divisibility condition and hence 311 and 281 (which do satisfy the divisibility condition) are the most likely cases. For the field with 311 elements, 168 and 169 are both quintic residues. Thus the argument that the triangle with vertices at 0, 1 and A cannot be interconnected by segments of the same color breaks down because of this unit difference in the quintic residues. The authors have not made a complete investigation of the case $p = 281$.

6. Relationship to Fermat's theorem and a certain trinomial congruence. Returning to the general case for a lower bound to $n(3^r)$, one requires that -1 be an rth power residue in a field of p elements. This requires a divisibility condition (already stated for 5), namely that r divides $\frac{1}{2}(p-1)$ when p is odd. One also requires that there be no solution of the trinomial congruence

$$1^r + x^r + y^r = 0$$

in the field. This congruence has been extensively studied by people interested in Fermat's theorem (**2**; **3**). Under the conditions assumed above a chromatic graph on p vertices with r colors can be so constructed that no monochromatic triangle appears.

Since an upper bound for $n(3^r)$ has been established by Theorem 6, it follows that when $p > n(3^r)$ and a chromatic graph has been constructed with r colors on the p field elements, a monochromatic triangle must appear. For such cases, then, the trinomial congruence is solvable.

The previous restriction that r divides $\frac{1}{2}(p-1)$ may be stated as r divides $p-1$ when r is odd. This may be interpreted as requiring the group of rth residues to have r cosets (including the original group of residues as one of its cosets). If r does not divide $p-1$, then in the field being an rth power is the same thing as being an mth power, where m is the greatest common divisor of r and $p-1$.

In view of the existence of the upper bound one obtains the result that there are only a finite number of finite fields for which the trinomial congruence is not solvable for given r (**3**).

References

1. L. E. Bush, *The William Lowell Putnam Mathematical Competition*, Amer. Math. Monthly, *60* (1953), 539–542.
2. H. S. Vandiver, *Fermat's Last Theorem*, Amer. Math. Monthly, *53* (1946), 555–578.
3. L. E. Dickson, *History of the theory of numbers*, Carnegie Institution of Washington, *2*, 763.

University of Texas *Harvard University*

REMARKS ON FINITE GROUPS DEFINED BY GENERATING RELATIONS

ROBERT FRUCHT

1. Introduction. After establishing a *duplication principle* (§2) which enables us to derive a group of order $2h$ with $k+1$ involutory generators from any group of order h with k generators, we shall prove the following *combination principle*:

Let $\mathfrak{F}$ and $\mathfrak{G}$ be two groups, each having k generators, such that the rth generators of $\mathfrak{F}$ and $\mathfrak{G}$ have relatively prime periods for each r from 1 to k. Then the direct product $\mathfrak{F} \times \mathfrak{G}$ can likewise be generated by k elements (instead of $2k$, as might be necessary without the condition imposed on the periods).

In the rest of the paper these two theorems are used as tools in the topological problem of finding symmetrical graphs of degree 3. Indeed, by a method that essentially goes back to Cayley and has already been used by the author in an earlier paper **(6)**, any group with three involutory generators gives rise to a graph of degree 3, and the "duplication principle" enables us to derive such a group from any group $\mathfrak{H}$ with two generators. The corresponding graph has $2n$ vertices, where n is the order of $\mathfrak{H}$, but it will be symmetrical only if $\mathfrak{H}$ satisfies a certain *symmetry condition* (§4). Some such groups can be combined by the "combination principle" to yield interesting examples of symmetrical graphs (§§5, 6).

I wish to take this opportunity to thank H. S. M. Coxeter for his criticism of the present paper in manuscript, and especially for adding some of the examples, and to thank R. M. Foster for telling me about the symmetrical graphs found by him.

2. The duplication principle. Let $\mathfrak{H}$ denote a finite abstract group whose generators $S_1, S_2, \ldots, S_k$ are subject to the sole defining relations

$$f_i(S_1, S_2, \ldots, S_k) = 1 \qquad (i = 1, 2, \ldots, m),$$

where $f_i(S_1, S_2, \ldots, S_k)$ denotes a product of powers of some or all of the generators. Let $\mathfrak{G}$ denote an abstract group whose generators $T_1, T_2, \ldots, T_{k+1}$ are subject to the sole defining relations

$$T_1^2 = T_2^2 = \ldots = T_{k+1}^2 = 1,$$
$$f_i(T_1T_2, T_2T_3, \ldots, T_k T_{k+1}) = 1 \qquad (i = 1, 2, \ldots, m).$$

Then *the order of $\mathfrak{G}$ is twice that of $\mathfrak{H}$.*

Received April 10, 1954.

Proof. The mapping

$$T_1T_2 \to S_1, \quad T_2T_3 \to S_2, \ldots, T_kT_{k+1} \to S_k$$

establishes an isomorphism between $\mathfrak{H}$ and a subgroup

$$\{T_1T_2, \quad T_2T_3, \ldots, T_kT_{k+1}\}$$

of $\mathfrak{G}$. This subgroup is of index 2 since it consists of all those elements of $\mathfrak{G}$ which are products of *even* numbers of T's, e.g.,

$$T_3T_1 = T_3T_2 \,.\, T_2T_1 = (T_2T_3)^{-1}(T_1T_2)^{-1}.$$

Example. Applying this principle to the cyclic group $\mathfrak{C}_n$, defined by

$$S^n = 1,$$

we obtain the dihedral group $\mathfrak{D}_{2n}$ (of order $2n$) in the form

$$T_1^2 = T_2^2 = (T_1T_2)^n = 1.$$

When $n = 2$, this is the four-group $\mathfrak{C}_2 \times \mathfrak{C}_2$. (Thus $\mathfrak{G}$ is sometimes, but not always, the direct product of $\mathfrak{H}$ and the group of order 2.)

More generally, whenever $\mathfrak{H}$ is Abelian and $k = 1$ or 2, $\mathfrak{G}$ is "generalized dihedral" in the sense of Miller (**9**, p. 168).

Remark. It is sometimes convenient to give $\mathfrak{H}$ an extra generator S_{k+1}, corresponding to $T_{k+1}T_1$, and to insert the extra relation

$$S_1S_2 \ldots S_kS_{k+1} = 1$$

(**10**, p. 26).

3. The combination principle. Let $\mathfrak{F}$ denote a finite group whose generators R_r $(r = 1, 2, \ldots, k)$ are subject to the sole defining relations

$$f_i(R_1, R_2, \ldots, R_k) = 1 \qquad (i = 1, 2, \ldots, m).$$

Let $\mathfrak{G}$ denote another group having the same number of generators $S_1, S_2, \ldots, S_k$, subject to the relations

$$g_i(S_1, S_2, \ldots, S_k) = 1 \qquad (i = 1, 2, \ldots, n).$$

Let a_r be the period of R_r (in $\mathfrak{F}$), and b_r that of S_r (in $\mathfrak{G}$). The direct product $\mathfrak{F} \times \mathfrak{G}$ can obviously be generated by $2k$ elements. *If, for each r, a_r and b_r are relatively prime, $\mathfrak{F} \times \mathfrak{G}$ is isomorphic with the group $\mathfrak{K}$ having only k generators, say U_r $(r = 1, 2, \ldots, k)$, subject to the defining relations*

$$f_i(U_1^{b_1}, U_2^{b_2}, \ldots, U_k^{b_k}) = 1 \qquad (i = 1, 2, \ldots, m),$$

$$g_i(U_1^{a_1}, U_2^{a_2}, \ldots, U_k^{a_k}) = 1 \qquad (i = 1, 2, \ldots, n),$$

$$U_r^{a_r}\, U_s^{b_s} = U_s^{b_s}\, U_r^{a_r} \qquad (r \neq s;\ r, s = 1, 2, \ldots, k).$$

Proof. $\mathfrak{K}$ contains a subgroup

$$\mathfrak{K}_F = \{U_1^{b_1}, U_2^{b_2}, \ldots, U_k^{b_k}\}$$

isomorphic with $\mathfrak{F}$, and a subgroup

$$\mathfrak{K}_G = \{U_1^{a_1}, U_2^{a_2}, \ldots, U_k^{a_k}\}$$

isomorphic with $\mathfrak{G}$. These have no element in common (except the identity), and each generator of $\mathfrak{K}_F$ commutes with each generator of $\mathfrak{K}_G$. Hence the $2k$ elements

$$U_r^{a_r}, U_r^{b_r} \qquad (r = 1, 2, \ldots, k)$$

will generate the direct product $\mathfrak{K}_F \times \mathfrak{K}_G$, isomorphic with $\mathfrak{F} \times \mathfrak{G}$. To verify that this subgroup $\mathfrak{K}_F \times \mathfrak{K}_G$ coincides with the whole group $\mathfrak{K}$, we express the generators U_r of the latter in terms of these particular powers, as follows. Since $(a_r, b_r) = 1$, we can find positive or negative integers α_r and β_r such that

$$\alpha_r a_r + \beta_r b_r = 1;$$

then

$$(U_r^{a_r})^{\alpha_r}(U_r^{b_r})^{\beta_r} = U_r \qquad (r = 1, 2, \ldots, k).$$

Example. Take $\mathfrak{F}$ to be the $\mathfrak{D}_{2n}$ defined by the generating relations

$$R_1^n = R_2^2 = (R_1R_2)^2 = 1,$$

and $\mathfrak{G}$ the same dihedral group, but with the generating relations

$$S_1^2 = S_2^n = (S_1S_2)^2 = 1.$$

Then if n is odd, the group $\mathfrak{K} = \{U_1, U_2\}$ defined by

$$U_1^{2n} = U_2^{2n} = (U_1^2U_2^n)^2 = (U_1^nU_2^2)^2 = 1,$$
$$U_1^nU_2^n = U_2^nU_1^n, \quad U_1^2U_2^2 = U_2^2U_1^2$$

will be isomorphic with the direct product $\mathfrak{D}_{2n} \times \mathfrak{D}_{2n}$.

Similarly, the direct product of two icosahedral groups,

$$R_1^5 = R_2^3 = (R_1R_2)^2 = 1$$

and

$$S_1^3 = S_2^5 = (S_1S_2)^2 = 1,$$

is

$$U_1^{15} = U_2^{15} = (U_1^3U_2^5)^2 = (U_1^5U_2^3)^2 = 1,$$
$$U_1^5U_2^5 = U_2^5U_1^5, \quad U_1^3U_2^3 = U_2^3U_1^3$$

(cf. **2**, p. 322).

Remark. When $k = 1$, this combination principle reduces to the well known fact that the direct product of two cyclic groups with relatively prime orders is again a cyclic group. Another special case, given by

$$b_2 = \ldots = b_k = 1,$$

is Theorem II of Carmichael (**1**, p. 167). An extension of our principle to more than two "factors" is obvious.

4. Application of the duplication principle to the construction of symmetrical graphs of degree three. Any group $\mathfrak{G}$ generated by three involutory elements has a graph of degree three for its "Cayley colour group"; e.g., the Abelian group $\mathfrak{C}_2 \times \mathfrak{C}_2 \times \mathfrak{C}_2$ is represented by the vertices and edges of a cube (**7**, p. 38). If $\mathfrak{G}$ admits an automorphism such that the three generators undergo a cyclic permutation, then the graph is symmetrical (**6**, p. 243). A great variety of symmetrical graphs can be derived from groups in this manner; the simplest that cannot is Petersen's graph (as R. M. Foster observed in a letter to the author).

A group $\mathfrak{G}$ whose generators T_1, T_2, T_3 are involutory can be derived by applying the duplication principle to any finite group $\mathfrak{H}$ with $k = 2$. Expressing $\mathfrak{H}$ in terms of three generators S_1, S_2, S_3 whose product is 1, we see that an automorphism of $\mathfrak{G}$ that cyclically permutes the T's arises from an automorphism of $\mathfrak{H}$ that cyclically permutes the S's. Accordingly, we say that a group $\mathfrak{H}$ with two generators, S_1 and S_2, fulfils the *symmetry condition* if it admits an automorphism such that the three elements S_1, S_2 and

$$S_3 = (S_1 S_2)^{-1} = S_2^{-1} S_1^{-1}$$

undergo a cyclic permutation.

Starting with any group $\mathfrak{H}$ of order n, that fulfils the symmetry condition, the duplication principle will yield a symmetrical graph of degree three having $2n$ vertices and $3n$ edges.

It should, however, be remarked that non-isomorphic groups of the same order can sometimes give rise to the same graph (as has been observed by Foster). For instance, Coxeter's graph $\{12\} + \{12/5\}$, formed by the 24 vertices and 36 edges of the hexagonal net $\{6, 3\}_{2,2}$ on a torus (**5**, pp. 437-440), arises from two distinct groups of order 12: the tetrahedral group $\mathfrak{T}_{12}$, defined by

4.1 $$S_1^3 = S_2^3 = (S_1^{-1}S_2)^2 = 1,$$

and the Abelian group

4.2 $$S_1^2 = S_2^2 = (S_1S_2)^{-2}, \quad S_1S_2 = S_2S_1$$

or

$$S_1^2 = S_2^2 = S_3^2, \quad S_1 S_2 S_3 = S_3 S_2 S_1 = 1,$$

which is the direct product $\mathfrak{C}_2 \times \mathfrak{C}_6$ or $\mathfrak{C}_2 \times \mathfrak{C}_2 \times \mathfrak{C}_3$.

It is clear that the graphs obtained by applying the duplication principle to suitable groups are always of the special kind which König calls "paar" (**8**, p. 170): the vertices of such a graph can be coloured alternately blue and red in such a manner that two vertices of the same colour are never joined by an edge. An important class of "paar" graphs are the "Levi graphs" of configurations (**5**, p. 413); e.g., the graph $\{12\} + \{12/5\}$, mentioned above, is the Levi graph for a configuration 12_3. Although symmetrical graphs that are not "paar" cannot be obtained by the duplication principle, they may still sometimes be obtainable as "Cayley colour groups"; e.g., Foster's graph of

girth 9 with 60 vertices (**11**, p. 459) represents the icosahedral group generated by the three permutations

$$(1\ 2)(3\ 5), \quad (1\ 3)(4\ 5), \quad (1\ 4)(2\ 5).$$

Incidentally, the same icosahedral group can also be generated by the two elements

$$S_1 = (1\ 2\ 3\ 4\ 5), \quad S_2 = (1\ 3\ 4\ 2\ 5),$$

which satisfy the symmetry condition with $S_3 = (1\ 4\ 2\ 3\ 5)$. Applying to this group $\mathfrak{H} = \{S_1, S_2\} = \{S_1, S_2, S_3\}$ the duplication principle, we obtain the extended icosahedral group $\mathfrak{C}_2 \times \mathfrak{H}$ generated by

$$T_1 = (1\ 2)(3\ 5)(6\ 7), \quad T_2 = (1\ 3)(4\ 5)(6\ 7), \quad T_3 = (1\ 4)(2\ 5)(6\ 7),$$

and this yields a graph of girth 10 with 120 vertices (also found by Foster).

5. Further examples of groups satisfying the symmetry condition.

5.1. Abelian groups (as well as some others such as the above-mentioned tetrahedral group) yield graphs whose girth is 6 (or less). Indeed, the relation $S_1S_2^{-1}S_1^{-1}S_2 = 1$, which holds in an Abelian group $\mathfrak{H} = \{S_1, S_2\}$, corresponds in $\mathfrak{G}$ to the relation

$$(T_1\, T_2\, T_3)^2 = 1,$$

which gives rise to hexagons in the graph.

In particular, the Abelian group

5.11 $$S_1^b = S_2^c = S_1^{-c}S_2^{-b}, \quad S_1S_2 = S_2S_1$$

or

$$S_1^b = S_2^c, \quad S_2^b = S_3^c, \quad S_3^b = S_1^c, \quad S_1S_2S_3 = S_3S_2S_1 = 1$$

yields the graph of vertices and edges of the regular map $\{6, 3\}_{b,c}$ on a torus (**5**, p. 421). This group, of order

$$n = b^2 + bc + c^2,$$

is $\mathfrak{C}_d \times \mathfrak{C}_{n/d}$, where $d = (b, c)$.

If $d = 1$, it is simply the cyclic group $\mathfrak{C}_n$. Since

$$(2b + c)^2 = 4n - 3c^2,$$

the possible values of n are

$$3, 7, 13, 19, 21, 31, \ldots,$$

for all of which (except the first) -3 is a quadratic residue. Thus $\mathfrak{C}_n$ yields a graph with $2n$ vertices whenever n is an odd prime not of the form $6m + 5$, or a square-free product of such primes. In this case the congruence

$$x^2 + x + 1 \equiv 0 \pmod{n}$$

or

$$(2x + 1)^2 \equiv -3 \pmod{4n}$$

has a solution x, and the same graph $\{6, 3\}_{b,c}$ arises from the same cyclic group $\mathfrak{C}_n$ in the form

$$S_1^n = 1, \quad S_1^x = S_2.$$

Since

$$S_1^x = S_2 \text{ and } S_2^x = S_1^{x^2} = S_1^{-x-1} = S_2^{-1}S_1^{-1},$$

the symmetry condition is fulfilled by the automorphism which transforms each element of $\mathfrak{C}_n$ into its xth power.

Setting $c = 0$ in 5.11, we obtain the direct product $\mathfrak{C}_b \times \mathfrak{C}_b$, defined by

$$S_1^b = S_2^b = 1, \quad S_1S_2 = S_2S_1$$

or

$$S_1^b = S_2^b = S_3^b = S_1S_2S_3 = S_3S_2S_1 = 1$$

(**3**, p. 97), which yields the graph $\{6, 3\}_{b,0}$ with $2b^2$ vertices.

Another instance, given by $c = b$, is the $\mathfrak{C}_b \times \mathfrak{C}_{3b}$ defined by

$$S_1^b = S_2^b = S_3^b, \quad S_1S_2S_3 = S_3S_2S_1 = 1,$$

which yields the graph $\{6, 3\}_{b,b}$ with $6b^2$ vertices.

5.2. A non-Abelian group of order $3b^2$, yielding the same graph $\{6, 3\}_{b,b}$, is

$$S_1^3 = S_2^3 = (S_1S_2)^3 = (S_1^{-1}S_2)^b = 1$$

or

$$S_1^3 = S_2^3 = S_3^3 = S_1S_2S_3 = (S_1^{-1}S_2)^b = (S_2^{-1}S_3)^b = (S_3^{-1}S_1)^b = 1$$

(**3**, pp. 99–100). When $b = 3$, these relations are equivalent to

$$S_1^3 = S_2^3 = 1, \quad S_1^{-1}S_2^{-1}S_1S_2 \,.\, S_i = S_i \,.\, S_1^{-1}S_2^{-1}S_1S_2 \qquad (i = 1, 2)$$

(**1**, p. 38, Ex. 26, 29) or

5.21 $$S_1^3 = S_2^3 = S_3^3 = S_1S_2S_3 = 1, \quad S_2S_1S_3 = S_3S_2S_1 = S_1S_3S_2.$$

Thus another generalization is

5.22 $$S_1^b = S_2^b = S_3^b = S_1S_2S_3 = 1, \quad S_2S_1S_3 = S_3S_2S_1 = S_1S_3S_2,$$

of order λb^2, where λ is the period of $S_3S_2S_1$, which is apparently b or $\frac{1}{2}b$ according as b is odd or even.

When $b = 4$, this is a group $\mathfrak{H}_{32}$ of order 32, defined by

5.23 $$S_1^4 = S_2^4 = (S_1S_2)^4 = 1, \quad S_1^2S_2 = S_2S_1^2, \quad S_1S_2^2 = S_2^2S_1,$$

which gives rise to the author's symmetrical graph of girth 8 with 64 vertices (**6**, p. 244 (3)).

Extending 5.21 another way, we obtain

$$S_1^3 = S_2^3 = S_3^3, \; S_1S_2S_3 = 1, \; S_2S_1S_3 = S_3S_2S_1 = S_1S_3S_2 = Z, \; Z^3 = 1,$$

of order 81, and

$$S_1^9 = S_2^9 = S_3^9 = S_1S_2S_3 = 1, \; S_2S_1S_3 = S_3S_2S_1 = S_1S_3S_2 = Z, \; Z^3 = 1,$$

of order 243. The former yields a new symmetrical graph of girth 12 with 162 vertices: The simplest known graph of girth 12.

5.3. The quaternion group $\mathfrak{Q}_8$, defined by

$$S_1^2 = S_2^2 = (S_1 S_2)^2$$

or

$$S_1^2 = S_2^2 = S_3^2, \quad S_1S_2S_3 = 1$$

gives rise to a graph of girth 6 with 16 vertices, the $\{8\} + \{8/3\}$ of Coxeter (**5**, p. 430).

Applying the combination principle to this and the $\mathfrak{C}_3$

$$R_1^3 = 1, \quad R_1 = R_2,$$

we obtain the direct product $\mathfrak{C}_3 \times \mathfrak{Q}_8$ in the form

$$U_1^{12} = 1, \quad U_1^4 = U_2^4,$$
$$U_1^6 = U_2^6 = (U_1^3U_2^3)^2,$$
$$U_1^3U_2^4 = U_2^4U_1^3, \quad U_1^4U_2^3 = U_2^3U_1^4,$$

which easily reduces to

$$U_1^2 = U_2^2 = (U_1U_2)^{-2}$$

or

$$U_1^{-2} = U_2^{-2} = U_3^2 = U_1U_2U_3,$$

the $\langle -2, -2, 2\rangle$ of Coxeter (**4**, pp. 367, 377). The duplication principle gives rise to a symmetrical graph of girth 8, with 48 vertices, found by Foster.

5.4. Applying the combination principle to the tetrahedral group $\mathfrak{T}_{12}$ and the four-group $\mathfrak{C}_2 \times \mathfrak{C}_2$, we obtain the group

$$S_1^6 = S_2^6 = S_3^6 = S_1S_2S_3 = (S_1^{-1}S_2)^2 = (S_2^{-1}S_3)^2 = (S_3^{-1}S_1)^2 = 1,$$

of order 48, which can be generated by permutations

$$(1\,2\,3)(5\,6)(7\,8), \quad (1\,3\,4)(7\,8)(9\,0), \quad (1\,4\,2)(9\,0)(5\,6),$$

and yields a graph of girth 8 with 96 vertices. Similarly

$$S_1^{12} = S_2^{12} = S_3^{12} = S_1S_2S_3 = (S_1^{-1}S_2)^2 = (S_2^{-1}S_3)^2 = (S_3^{-1}S_1)^2 = 1,$$

of order 96 (**3**, p. 101, footnote), yields a new graph of girth 8 with 192 vertices.

5.5. The group

5.51 $$S_1^{2n} = S_2^{2n} = (S_1S_2)^{2n} = 1, \quad S_1^2S_2S_1^2 = S_2, \quad S_2^2S_1S_2^2 = S_1$$

or

$$S_1^{2n} = S_2^{2n} = S_3^{2n} = S_1S_2S_3 = 1, \quad S_1^2S_2S_1^2 = S_2, \quad S_2^2S_3S_2^2 = S_3, \quad S_3^2S_1S_3^2 = S_1,$$

of order $4n^3$, yields a graph with $8n^3$ vertices, which has girth 10 if $n > 3$.

Thus a graph of girth 10 with 216 vertices comes from the group of order 108 generated by

$$(1\,2)(4\,5\,6)(8\,9),\quad (2\,3)(4\,5)(7\,8\,9),\quad (1\,2\,3)(5\,6)(7\,8).$$

The above relations are easily seen to imply

$$S_1^2S_2^2 = S_2^2S_1^2,\quad S_2^2S_3^2 = S_3^2S_2^2,\quad S_3^2S_1^2 = S_1^2S_3^2,$$

so that the subgroup $\{S_1^2, S_2^2, S_3^2\}$, of index 4, is simply $\mathfrak{C}_n \times \mathfrak{C}_n \times \mathfrak{C}_n$. If n is even, this subgroup has a factor group of order $\frac{1}{2}n^3$, given by inserting the extra relation

$$S_1^nS_2^nS_3^n = 1 \text{ or } (S_1^2S_2^2S_3^2)^{\frac{1}{2}n} = 1 \text{ or } (S_3S_2S_1)^{\frac{1}{2}n} = 1.$$

Hence the whole group has a factor group of order $2n^3$ (when n is even), given by inserting the extra relation

$$(S_1^{-1}S_2^{-1}S_1S_2)^{\frac{1}{2}n} = 1 \text{ or } (S_3S_2S_1)^{\frac{1}{2}n} = 1.$$

This yields a graph with $4n^3$ vertices (n even), whose girth is 10 if $n > 4$. Thus a graph of girth 10 with 256 vertices comes from the group of order 128 defined by

$$S_1^8 = S_2^8 = (S_1S_2)^8 = (S_1^{-1}S_2^{-1}S_1S_2)^2 = 1,\ S_1^2S_2S_1^2 = S_2,\ S_2^2S_1S_2^2 = S_1.$$

5.6. Similarly, the group

5.61 $$S_1^{2n} = S_2^{2n} = (S_1S_2)^{2n} = 1,\quad S_1^2S_2 = S_2S_1^2,\quad S_1S_2^2 = S_2^2S_1$$

or

$$S_1^{2n} = S_2^{2n} = S_3^{2n} = S_1S_2S_3 = 1,\ S_1^2S_2 = S_2S_1^2,\ S_2^2S_3 = S_3S_2^2,\ S_3^2S_1 = S_1S_3^2,$$

of order $4n^3$, yields a graph with $8n^3$ vertices whose girth is 10 if $n > 3$; and if n is even, it has a factor group of order $2n^3$, given by inserting the extra relation

$$(S_1^{-1}S_2^{-1}S_1S_2)^{\frac{1}{2}n} = 1 \text{ or } (S_3S_2S_1)^{\frac{1}{2}n} = 1,$$

which yields a graph with $4n^3$ vertices.

When $n = 2$, 5.51 and 5.61 both reduce to 5.23. But when $n > 2$, the two graphs are distinct, as R. M. Foster found by examining them in detail.

5.7. A graph of girth 12, with 288 vertices, is yielded by the group

$$S_1^6 = S_2^6 = S_3^6 = S_1S_2S_3 = (S_3S_2S_1)^2 = S_3^3S_2^3S_1^3 = 1,$$
$$S_1^3S_2 = S_2S_1^3,\quad S_2^3S_3 = S_3S_2^3,\quad S_3^3S_1 = S_1S_3^3$$

of order 144, which is a subgroup of $\mathfrak{C}_6 \times \mathfrak{C}_6 \times \mathfrak{T}_{12}$.

5.8. A graph of girth 8, with 336 vertices, results from the well-known simple group of order 168 in the form

$$S_1^4 = S_2^4 = (S_1^{-1}S_2)^3 = (S_1^2S_2)^3 = 1$$

or

$$S_1^4 = S_2^4 = S_3^4 = S_1S_2S_3 = (S_1^{-1}S_2)^3 = (S_2^{-1}S_3)^3 = (S_3^{-1}S_1)^3 = 1$$

(**3**, p. 96). As permutations of degree seven, we have

$$S_1 = (1\,2)(4\,6\,5\,7),\quad S_2 = (3\,4)(6\,2\,1\,7),\quad S_3 = (5\,6)(2\,4\,3\,7).$$

5.9. The author's one-regular graph of girth 12, with 432 vertices (**6**, p. 246), comes from Hesse's collineation group of order 216 (**9**, p. 239) in the form

$$S_1^6 = S_2^6 = S_3^6 = S_1S_2S_3 = 1,\ S_1^2S_2S_1^2 = S_2^2,\ S_2^2S_3S_2^2 = S_3^2,\ S_3^2S_1S_3^2 = S_1^2.$$

As permutations of degree 9, we have

$$S_1 = (4\,8)(1\,6\,2\,3\,9\,5),\quad S_2 = (5\,9)(3\,4\,1\,2\,7\,6),\quad S_3 = (6\,7)(2\,5\,3\,1\,8\,4).$$

6. Concluding remarks. In each of the foregoing examples the duplication principle would allow us to indicate also generating relations for a group $\mathfrak{G}$ whose order is twice that of the given group $\mathfrak{H}$. For instance, the group 5.22 yields

$$T_1^2 = T_2^2 = T_3^2 = (T_2T_3)^b = (T_3T_1)^b = (T_1T_2)^b = 1,$$
$$(T_1T_2T_3)^2 = (T_2T_3T_1)^2 = (T_3T_1T_2)^2,$$

of order $2\lambda b^2$ (where $\lambda = b$ or $\frac{1}{2}b$ according as b is odd or even).

Again, 5.51 and 5.61 yield

$$T_1^2 = T_2^2 = T_3^2 = (T_2T_3)^{2n} = (T_3T_1)^{2n} = (T_1T_2)^{2n} = 1$$

with the extra relations

$$T_1T_2T_1T_3T_1 = T_2T_3T_2T_1T_2 = T_3T_1T_3T_2T_3$$

and

$$[T_1(T_2T_3)^2]^2 = [T_2(T_3T_1)^2]^2 = [T_3(T_1T_2)^2]^2 = 1,$$

respectively. When $n = 3$, the former has a factor group of order 108, given by the extra relation $(T_1T_2T_3)^3 = 1$; this yields Foster's graph of girth 9 with 108 vertices.

If the conditions for the combination principle are fulfilled, its application yields further symmetrical graphs corresponding to such direct products as

$$\mathfrak{C}_7 \times \mathfrak{Q}_8,\quad (\mathfrak{C}_3 \times \mathfrak{C}_3) \times \mathfrak{Q}_8,\quad \mathfrak{C}_7 \times \mathfrak{T}_{12},$$
$$\mathfrak{C}_3 \times \mathfrak{H}_{32},\quad \mathfrak{Q}_8 \times \mathfrak{T}_{12},\quad \mathfrak{Q}_8 \times \mathfrak{C}_{13},$$

etc. When generating relations for each "factor" are known, the combination principle yields also generating relations for the direct product, as in §5.3 for $\mathfrak{C}_3 \times \mathfrak{Q}_8$, and in §5.4 for $(\mathfrak{C}_2 \times \mathfrak{C}_2) \times \mathfrak{T}_{12}$.

References

1. R. D. Carmichael, *Introduction to the theory of groups of finite order* (Boston), 1937.
2. H. S. M. Coxeter, *Abstract definitions for the symmetry groups of the regular polytopes, in terms of two generators*. Part II: *The rotation groups*, Proc. Camb. Philos. Soc., *33* (1937), 315–324.
3. ——— *The abstract groups* $G^{m,n,p}$, Trans. Amer. Math. Soc., *45* (1939), 73–150.
4. ——— *The binary polyhedral groups, and other generalizations of the quaternion group*, Duke Math. J., 7 (1940), 367–379.
5. ——— *Self-dual configurations and regular graphs*, Bull. Amer. Math. Soc., *56* (1950), 413–455.
6. R. Frucht, *A one-regular graph of degree three*, Can. J. Math., *4* (1952), 240–247.
7. A. B. Kempe, *A memoir on the theory of mathematical form*, Phil. Trans. Royal Soc. A, *177* (1886), 1–70.
8. D. König, *Theorie der endlichen und unendlichen Graphen* (Leipzig, 1936).
9. G. A. Miller, H. F. Blichfeldt, and L. E. Dickson, *Theory and applications of finite groups* (New York, 1916).
10. W. Threlfall, *Gruppenbilder*, Abh. Math.-Phys. Klasse Sächs. Akad. Wiss., *41.6* (1932), 1–59.
11. W. T. Tutte, *A family of cubical graphs*, Proc. Cambridge Philos. Soc., *43* (1947), 459–474.

Technical University "Santa Maria"
Valparaiso, Chile

ON LAVES' GRAPH OF GIRTH TEN

H. S. M. COXETER

1. Introduction. This note shows how a certain infinite graph of degree three, discovered by Laves in connection with crystal structure, can be inscribed (in sixteen ways, all alike) in an infinite regular skew polyhedron which has square faces, six at each vertex. One-eighth of the vertices of the polyhedron are vertices of the graph, and the three edges of the graph that meet at such a vertex are diagonals of alternate squares. Thus either diagonal of any face of the polyhedron can serve as an edge, and the whole graph can then be completed in a unique manner.

The same graph is also derived, by the method of Frucht's paper **(4)**, from the abstract group

$$S_1^2S_2S_1^2 = S_2, \quad S_2^2S_1S_2^2 = S_1.$$

2. A regular skew polyhedron. In 1926, J. F. Petrie discovered an infinite regular skew polyhedron which can be derived from the simple honeycomb of cubes $\{4, 3, 4\}$ by taking all the vertices, all the edges, and half the squares **(1**, pp. 33–35; **2**, p. 242; **6**, p. 55, Fig. 4**)**. This new regular polyhedron was named $\{4, 6\,|\,4\}$ because it has square faces, six at each vertex, and square holes corresponding to the missing faces of the cubic honeycomb.

The vertex figure of a cube $\{4, 3\}$ is an equilateral triangle $\{3\}$ **(3**, p. 16**)**; the vertex figure of the cubic honeycomb $\{4, 3, 4\}$ is an octahedron $\{3, 4\}$ **(3**, p. 68**)**; and the vertex figure of the skew polyhedron $\{4, 6\,|\,4\}$ is a skew hexagon which is a Petrie polygon of the octahedron **(3**, pp. 24–25**)**. Since a given octahedron has four Petrie polygons, the faces of the skew polyhedron can be selected from those of a given cubic honeycomb in four ways. Taking the vertices to be all the points whose Cartesian coordinates are integers, one way is select the squares

$$(0, 0, 0)(0, 1, 0)(0, 1, -1)(0, 0, -1)$$
$$(1, 0, 0)(1, 1, 0)(1, 1, -1)(1, 0, -1)$$

and all others that can be derived from these two by permuting the three coordinates and adding fixed even numbers to them.

3. The infinite group $\{R, S\}$. In saying that the skew polyhedron is "regular," we mean that it possesses two special symmetry operations: R, cyclically permuting the vertices of one face (say the former of those mentioned above), and S, cyclically permuting the faces at one vertex (say the

Received June 8, 1954.

origin). These operations generate an infinite group having the elegant abstract definition

$$R^4 = S^6 = (RS)^2 = (RS^{-1})^4 = 1$$

(**1**, pp. 35, 48). Since they are rotatory reflections (**1**, p. 37), their squares are pure rotations: R^2, of period 2 about the "axis" of the square

$$(0, 0, 0)(0, 1, 0)(0, 1, -1)(0, 0, -1),$$

and S^2, of period 3 about the line $x = y = z$. In fact, R^2 is the transformation

$$x' = x, \quad y' = 1 - y, \quad z' = -1 - z,$$

while S^2 is the cyclic permutation

$$x' = z, \quad y' = x, \quad z' = y.$$

4. The subgroups $\{T_1, T_2, T_3\}$ and $\{S_1, S_2, S_3\}$. It is convenient to let T_1, T_2, T_3 denote R^2 and its transforms by S^2 and S^{-2}, namely

$$\begin{array}{llll} T_1: & x' = x, & y' = 1 - y, & z' = -1 - z, \\ T_2: & x' = -1 - x, & y' = y, & z' = 1 - z, \\ T_3: & x' = 1 - x, & y' = -1 - y, & z' = z. \end{array}$$

We use S_1, S_2, S_3 to denote the products T_2T_3, T_3T_1, T_1T_2, namely

$$\begin{array}{llll} S_1: & x' = 2 + x, & y' = -1 - y, & z' = 1 - z, \\ S_2: & x' = 1 - x, & y' = 2 + y, & z' = -1 - z, \\ S_3: & x' = -1 - x, & y' = 1 - y, & z' = 2 + z. \end{array}$$

Thus S_1^2, S_2^2, S_3^2 increase the respective coordinates by 4, and $S_1T_2S_3$ increases each of them by 2.

The half-turns T_i can be expressed in an obvious way as products of pairs of reflections A_i and B_i, namely

$$T_1 = A_3B_2, \quad T_2 = A_1B_3, \quad T_3 = A_2B_1,$$

where

$$\begin{array}{lllll} A_1 & \text{is} & x' = -1 - x, & y' = y, & z' = z, \\ B_1 & \text{is} & x' = 1 - x, & y' = y, & z' = z, \\ A_2 & \text{is} & x' = x, & y' = -1 - y, & z' = z, \end{array}$$

and so on. Thus the A's and B's are reflections in the opposite faces

$$x = \mp \tfrac{1}{2}, \quad y = \mp \tfrac{1}{2}, \quad z = \mp \tfrac{1}{2}$$

of the cube $(\pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2})$. They generate the direct product of three infinite groups

$$A_i^2 = B_i^2 = 1 \qquad (i = 1, 2, 3),$$

since both A_i and B_i commute with both A_j and B_j whenever $i \neq j$. The

products A_iB_i increase the respective coordinates by 2. We easily verify that

$$(A_iB_i)^2 = S_i^2$$

and

$$A_1B_1A_2B_2A_3B_3 = A_1A_2A_3B_1B_2B_3 = T_2T_3T_2T_1T_2 = S_1T_2S_3.$$

5. An infinite graph. Since the three T's are half-turns about three edges of the above-mentioned cube, which are the axes of three faces of $\{4, 6 \mid 4\}$ (see Fig. 1), they transform the origin into the three points

$$(0, 1, -1), \quad (-1, 0, 1), \quad (1, -1, 0)$$

which are the opposite vertices of these squares. Thus the group $\{T_1, T_2, T_3\}$ transforms the origin into the points

$$\begin{array}{llll} (0, 0, 0), & (1, 2, 3), & (2, 3, 1), & (3, 1, 2) \\ (2, 2, 2), & (3, 0, 1), & (0, 1, 3), & (1, 3, 0) \end{array} \qquad \text{(mod 4)},$$

which are derived from $(0, 0, 0)$ and $(1, 2, 3)$ by cyclically permuting the coordinates, adding 2 to all of them, and adding arbitrary multiples of 4 to any of them.

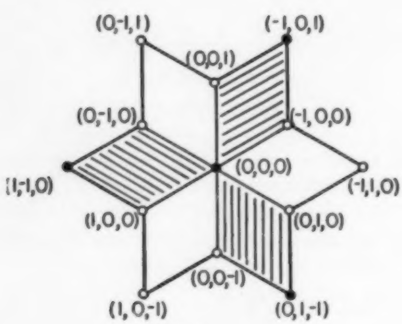

Figure 1

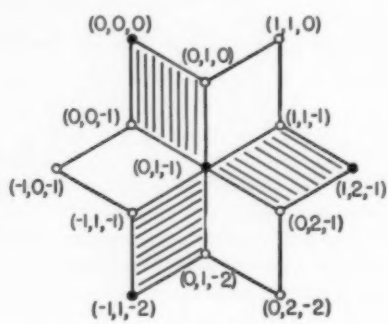

Figure 2

In this manner we pick out one-eighth of the vertices of $\{4, 6 \mid 4\}$, occurring at the ends of diagonals of one-eighth of the squares (shaded in Figs. 1 and 2). These diagonals are the edges of a graph of degree three which is the Cayley colour group for $\{T_1, T_2, T_3\}$. It is remarkable that the three edges at a vertex are coplanar and meet at angles of 120°, like those of the plane tessellation of hexagons, $\{6, 3\}$.

The smallest circuit in this graph is a skew decagon such as

$$(0, 0, 0)\ (0, 1, -1)\ (1, 2, -1)\ (1, 3, 0)\ (2, 3, 1)\ (2, 2, 2)\ (1, 2, 3)\ (0, 1, 3)$$
$$(-1, 1, 2)\ (-1, 0, 1).$$

In other words, the graph is of *girth* ten (**9**, p. 459). This property corresponds to the fact that, apart from

$$T_1^2 = T_2^2 = T_3^2 = 1,$$

the simplest relations satisfied by the T's are

$$T_1T_2T_1T_3T_1 = T_2T_3T_2T_1T_2 = T_3T_1T_3T_2T_3$$

(i.e., $S_3T_1S_2 = S_1T_2S_3 = S_2T_3S_1$, or $S_i^2T_i = T_iS_i^2$), each involving ten T's.

In terms of the S's alone, which evidently satisfy

$$S_1S_2S_3 = 1,$$

these relations become

$$S_1^2S_2S_1^2 = S_2, \quad S_2^2S_3S_2^2 = S_3, \quad S_3^2S_1S_3^2 = S_1$$

(any two of which imply the remaining one). The first implies

$$S_1^2S_2^2 = S_1^2S_2.S_1^2S_2S_1^2 = S_1^2S_2S_1^2.S_2S_1^2 = S_2^2S_1^2;$$

thus the squares of the S's all commute with one another. By repeated application of

$$S_1S_2^2 = S_2^{-2}S_1, \quad S_2S_3^2 = S_3^{-2}S_2, \quad S_3S_1^2 = S_1^{-2}S_3,$$
$$S_1S_3^2 = S_3^{-2}S_1, \quad S_2S_1^2 = S_1^{-2}S_2, \quad S_3S_2^2 = S_2^{-2}S_3,$$

we can transform any "word" into the standard form

$$S_1^xS_2^yS_3^z$$

where at most one of x, y, z is odd; e.g.,

$$S_1S_3 = S_1^2S_1^{-1}S_3^{-1}S_3^2 = S_1^2S_2S_3^2.$$

Thus the Abelian subgroup $\mathfrak{S} = \{S_1^2, S_2^2, S_3^2\}$, having the four cosets

$$\mathfrak{S}, \ \mathfrak{S}S_1, \ \mathfrak{S}S_2, \ \mathfrak{S}S_3,$$

is of index 4 in $\{S_1, S_2, S_3\}$.

6. A finite graph. The abstract group $\{S_1, S_2, S_3\}$ has a finite factor group

$$S_1^{2l} = S_2^{2m} = S_3^{2n} = S_1S_2S_3 = 1,$$
$$S_1^2S_2S_1^2 = S_2, \quad S_2^2S_3S_2^2 = S_3, \quad S_3^2S_1S_3^2 = S_1$$

or

$$S_1^{2l} = S_2^{2m} = (S_1S_2)^{2n} = 1, \quad S_1^2S_2S_1^2 = S_2, \quad S_2^2S_1S_2^2 = S_1,$$

whose order is $4lmn$ (since it has a subgroup $\mathfrak{C}_l \times \mathfrak{C}_m \times \mathfrak{C}_n$ of index 4). The case $l = m = n$ has been described by Frucht (**4**, 5.51).

Geometrically, since the squared S's in the infinite group are translations, the effect of giving them definite periods l, m, n is to identify points whose three coordinates differ by multiples of $4l, 4m, 4n$, respectively. In other words, the infinite space is reduced to a three-dimensional torus. Since the order of

the finite group is $4lmn$, the corresponding graph has $8lmn$ vertices and $12lmn$ edges.

The girth is still 10, provided l, m, n are all greater than 2. But if $l = 2$, the relation $S_1^4 = 1$ or $(T_2 T_3)^4 = 1$, involving 8 letters, yields an 8-circuit. In the special case $l = m = n = 2$, every edge belongs to just two 8-circuits (e.g., the "T_2" edge belongs to $(T_1T_2)^4 = 1$ and $(T_2T_3)^4 = 1$); therefore the graph **(4, 5.23)** can be embedded in a surface to form a regular map of 24 octagons.

7. A thin packing of spheres. Returning to the infinite graph (with l, m, n unrestricted), we note that something essentially equivalent to it was discovered in 1932 by Laves **(7**, p. 10**)**. When the graph is derived from the $\{4, 6 | 4\}$ whose vertices have integral coordinates, its edges are of length $\sqrt{2}$. Hence spheres of diameter $\sqrt{2}$, drawn around all the vertices, will touch one another at the mid-points of the edges. This arrangement of spheres is the figure described by Laves. His Fig. 7 shows, in a very striking manner, some of the rings of ten spheres corresponding to 10-circuits in the graph.

Since there are eight spheres for each cube of edge 4, the packing density is

$$8 \cdot \frac{4}{3} \pi(\sqrt{\tfrac{1}{2}})^3/4^3 = \frac{\pi}{12\sqrt{2}} = 0.18512\ldots,$$

namely one-quarter that of the cubic or hexagonal close-packing. Nevertheless, this is not the thinnest possible packing of equal spheres. By the simple but ingenious device of replacing each sphere by a cluster of three smaller ones **(5**, pp. 448–450; **8**, p. 484**)**, Heesch and Laves derived a still thinner packing, with density only

$$\pi\sqrt{2}\left(\sqrt{3} - \frac{3}{2}\right)^3 = 0.055515\ldots.$$

8. A correction. I take this opportunity to correct an unfortunate error in my paper on *Regular skew polyhedra* **(1**, pp. 54, 55, 61**)**. The regular map shown in Fig. XV is not $\{4, 7 | 3\}$ but $\{4, 6 | 3\}$. Moreover, Table II should be supplemented by three further entries:

Polyhedron	f	e	v	p	$\mathfrak{G}$	g
$\{4, 6 \,|, 2\}$	12	24	8	3	$S_4 \times S_2$	48
$\{5, 6 \,|, 2\}$	24	60	20	9	$A_5 \times S_2$	120
$\{3, 11 \,|, 4\}$	2024	3036	552	231	$LF(2, 23)$	6072

Note added in proof. When applied to the four-dimensional polyhedron $\{4, 6 | 3\}$ **(1**, pp. 45, 55**)**, the procedure of §§ 1 and 5 (using diagonals of alternate squares at a vertex) yields a graph of girth five having 20 vertices and 30

edges (one diagonal of each square); this is the same as the graph formed by the vertices and edges of the regular dodecahedron $\{5, 3\}$. When applied to $\{4, 6|, 2\}$, it yields the complete 4-point (i.e., the vertices and edges of the tetrahedron $\{3, 3\}$). When applied to $\{4, 6|, 3\}$ (**1**, p. 60), it yields a graph of girth six having 24 vertices and 36 edges (diagonals of 36 of the 84 squares). This graph, denoted by $\{12\} + \{12/5\}$ (**4**, 4.1), consists of the vertices and edges of the map $\{6, 3\}_{2,2}$ on a torus, and is the Cayley colour group for the octahedral group S_4 generated by the three transpositions $T_i = (i\ 4)$ $(i = 1, 2, 3)$.

References

1. H. S. M. Coxeter, *Regular skew polyhedra in three and four dimensions, and their topological analogues*, Proc. London Math. Soc. (2), *43* (1937), 33–62.
2. ———, *The regular sponges, or skew polyhedra*, Scripta Math., *6* (1939), 240–244.
3. ———, *Regular polytopes* (London, 1948; New York, 1949).
4. R. Frucht, *Remarks on finite groups defined by generating relations*, Can. J. Math., *7* (1955), 8–17.
5. H. Heesch and F. Laves, *Über dünne Kugelpackungen*, Z. Kristallogr., *85* (1933), 443–453.
6. H. Jensen, *De seks uendelige regulaere polyedre*, Mat. Tidsskr. (A), 1950, 53–60.
7. F. Laves, *Zur Klassifikation der Silikate*, Z. Kristallogr., *82* (1932), 1–14.
8. S. Melmore, *The densest and the least dense packings of equal spheres*, Mineralogical Magazine, *28* (1949), 479–485.
9. W. T. Tutte, *A family of cubical graphs*, Proc. Cambridge Philos. Soc., *43* (1947), 459–474.

University of Toronto

A CLASS OF PERMUTATION GROUPS OF PRIME DEGREE

K. D. FRYER

1. Introduction. In **(1)**, using the theory of group representations, Brauer studied groups $\mathfrak{G}$ of finite order g containing elements A of prime period p which commute only with their own powers A^i. If $\mathfrak{P}$ is a p-Sylow subgroup of $\mathfrak{G}$, the normalizer $\mathfrak{N} = \mathfrak{N}(\mathfrak{P})$ of $\mathfrak{P}$ can be generated by A and another element B such that

$$A^p = 1, \quad B^q = 1, \quad B^{-1}AB = A^{\gamma^t}, \tag{1.1}$$

where γ is a primitive root (mod p), and t and q are positive integers such that

$$tq = p - 1. \tag{1.2}$$

For A, B satisfying 1.1, the group $\{A, B\}$ is of order pq. Carmichael **(3)** points out that if $t = 1$, the group $\{A, B\}$ is the so-called *metacyclic* group of order $p(p-1)$, simply isomorphic with a doubly transitive group of degree p. Such a group contains a dihedral subgroup and is complete.

The permutations

$$A = (0\,1\,\ldots\,6), \quad B = (1\,2\,4)(3\,6\,5)$$

in $GF(7)$ satisfy 1.1 with $p = 7$, $q = 3$, $t = 2$, $\gamma = 3$, and generate a group of order 21. If now a third permutation $C_1 = (2\,4)(5\,6)$ is added, the group $\{A, B, C_1\}$ generated is the linear fractional group $LF(2, 7)$ of order 168. If, instead of C_1, we add $C_2 = (2\,4)(3\,5)$, the group $\{A, B, C_2\}$ is the alternating group $\mathfrak{A}_7$, while for $C_3 = (1\,6)(2\,3)(4\,5)$, $\{A, B, C_3\}$ is the symmetric group $\mathfrak{S}_7$.

The elements C_1, C_2, C_3 stand in the same relation to B as does B to A; that is, they satisfy the relations

$$B^q = 1, \quad C^r = 1, \quad C^{-1}BC = B^{\delta^s}. \tag{1.3}$$

where δ is a primitive root (mod q), and where s and r are positive integers such that

$$sr = q - 1. \tag{1.4}$$

For these elements, $q = 3$, $r = 2$, $s = 1$, $\delta = 2$.

The permutations

$$A = (0\,1\,2\,\ldots\,10), \quad B = (1\,4\,5\,9\,3)(2\,8\,10\,7\,6)$$

in $GF(11)$ satisfy 1.1 with $p = 11$, $q = 5$, $t = 2$, $\gamma = 2$.

Received June 2, 1953; in revised form March 30, 1954. This paper comprises the substance of the author's thesis, written under the direction of Professors R. Brauer and R. G. Stanton and accepted for a Ph.D. degree at the University of Toronto in June, 1952.

The additional permutations

$$C_1 = (2\ 8)(6\ 10)(3\ 4)(5\ 9), \qquad C_2 = (2\ 7)(8\ 10)(3\ 4)(5\ 9),$$
$$C_3 = (1\ 10)(2\ 5)(3\ 7)(4\ 8)(6\ 9), \quad C_4 = (2\ 10\ 8\ 6)(3\ 9\ 4\ 5)$$

yield respectively the following groups $\{A, B, C\}$:

$$LF(2, 11), \quad \mathfrak{A}_{11}, \quad \mathfrak{S}_{11}, \quad \mathfrak{M}_{11}.$$

The permutations C_i satisfy relations 1.3. For C_1, C_2, C_3, we have $q = 5$, $r = 2$, $\delta = 2$, $s = 2$; for C_4, $q = 5$, $r = 4$, $\delta = 2$, $s = 1$.

Finally, for $p = 23$, $q = 11$, $t = 2$, $\gamma = 5$, the permutations

$$A = (0\ 1\ 2 \ldots 22),$$
$$B = (1\ 2\ 4\ 8\ 16\ 9\ 18\ 13\ 3\ 6\ 12)(5\ 10\ 20\ 17\ 11\ 22\ 21\ 19\ 15\ 7\ 14)$$

in $GF(23)$ satisfy 1.1, and the additional permutations

$$C_1 = (2\ 16\ 9\ 6\ 8)(4\ 3\ 12\ 13\ 18)(7\ 17\ 10\ 11\ 22)(14\ 19\ 21\ 20\ 15),$$
$$C_2 = (2\ 12)(4\ 6)(8\ 3)(16\ 13)(9\ 18)(7\ 19)(14\ 21)(5\ 22)(10\ 11)(17\ 20),$$
$$C_3 = (1\ 22)(2\ 11)(3\ 15)(4\ 17)(5\ 9)(6\ 19)(7\ 13)(8\ 20)(10\ 16)(12\ 21)(14\ 18)$$

yield[1] in turn $\mathfrak{M}_{23}$, $\mathfrak{A}_{23}$, $\mathfrak{S}_{23}$. For C_1, $q = 11$, $r = 5$, $s = 2$, $\delta = 2$; for C_2 and C_3, $q = 11$, $r = 2$, $s = 5$, $\delta = 2$.

We have thus been led to a study of groups $\mathfrak{H} = \{A, B, C\}$ whose generators satisfy the following abstract relations:

1.5 $$A^p = B^q = C^r = 1, \quad B^{-1}AB = A^j, \quad C^{-1}BC = B^l,$$

where p and q are primes, $p = 2q + 1$, r is an arbitrary divisor of $q - 1$, $q = sr + 1$, j belongs to q modulo p, and l belongs to r modulo q.

We prove the following

THEOREM. *The groups $\{A, B, C\}$ described above fall into two classes: groups consisting entirely of even permutations (case* I*), and groups of even and odd permutations (case* II*). The groups in case* II *are the symmetric groups $\mathfrak{S}_p$. The groups in case* I *with r even and $\beta = q$ are the alternating groups $\mathfrak{A}_p$.* (The explanation of the symbol β is given in §2).

It is demonstrated by tables in the author's thesis, to be found in the library at the University of Toronto, that for $p < 59$, the groups $\mathfrak{H} = \{A, B, C\}$ in case I are the alternating groups $\mathfrak{A}_p$ with the exception of the four groups mentioned above; that is, $LF(2, 7)$, $LF(2, 11)$, $\mathfrak{M}_{11}$ and $\mathfrak{M}_{23}$. The same result also holds for $p = 83$.

We prove that the groups in case I are all simple. If it could be proved that there is an infinite number of prime pairs p, q with $p = 2q + 1$, we would

[1]Carmichael **(3**, p. 288**)** indicates that the groups $\mathfrak{M}_{11}$ and $\mathfrak{M}_{23}$ are generated by permutations similar to the A and C specifically mentioned above. It can be shown that the corresponding B in each case is expressible in terms of A and C. For further information on these groups, see **(6)**.

have here an infinite class of simple groups containing in particular the two Mathieu groups $\mathfrak{M}_{11}$ and $\mathfrak{M}_{23}$.

2. The permutation representation of the generators. We proceed first to develop permutation representations for the generators A, B, C which satisfy the relations 1.5. It must be stressed that in the following we have $p = 2q + 1$ for p and q primes, and there exists a natural mapping of the integers mod q on the quadratic residues $\neq 0$ mod p and on the quadratic non-residues mod p. This fact enables us to prove

THEOREM 2.1. *The group generators A, B, and C of* §1 *may be represented by the following permutations in $GF(p)$:*

$$\begin{aligned} &A\colon\ x' = x + 1,\\ &B\colon\ x' = g^2 x, \qquad\qquad (2.1)\\ &C\colon\ x' = \tfrac{1}{2}(1+\epsilon)\, g^{\alpha(1-l)}x^l + \tfrac{1}{2}(1-\epsilon)\, g^{\beta(1-l)}x^l \end{aligned}$$

where $\epsilon = \pm 1$ according as x is a quadratic residue or non-residue mod p, *g is that primitive root* mod p *for which $g^2 = j$, and α and β are arbitrary fixed even and odd numbers, respectively.*

Our procedure is briefly as follows: we notice that A and B generate the metacyclic group $\{A, B\}$, and obtain representations on p symbols for A and B:

2.2 $$A\colon\ x' = f(x), \quad B\colon\ x' = g(x), \qquad x = 0, 1, \ldots p - 1.$$

B and C generate another metacyclic group, $\{B, C\}$, and we obtain representations on q symbols for B and C:

2.3 $$B\colon\ y' = g'(y), \quad C\colon\ y' = h'(y), \qquad y = 0, 1, \ldots, q - 1.$$

We then combine the representations 2.2 and 2.3, identifying the two representations for B, and obtain a representation for A, B, and C on p symbols:

2.4 $$A\colon\ x' = f(x), \quad B\colon\ x' = g(x), \quad C\colon\ x' = h(x) \qquad x = 0, 1, \ldots, p - 1.$$

Consider the relations involving A and B, namely

$$A^p = B^q = 1, \quad B^{-1}AB = A^j.$$

These imply

$$B^{-1}A^xB = A^{xj}$$

and

$$B^{-y}A^xB^y = A^{xj^y}.$$

Hence $A^xB^y = B^yA^{xj^y}$, and the pq elements A^xB^y correspond to the pq elements

$$B^yA^z, \qquad\qquad z = xj^y,$$

in some order. For as x ranges from 0 to $p - 1$, $x \to z = xj^y$ is a permutation

of the numbers $0, 1, \ldots, p-1$, for a fixed value of y. So the elements A and B generate the metacyclic group of order pq, the so-called *Cauchy* group **(2)**.

Let A be represented as the cycle $(0\,1\,2 \ldots p-1)$, that is, the transformation $x' = x + 1$ in $GF(p)$. Let B be the permutation $x' = g(x)$. Since $AB = BA^j$, we have

$$g(x+1) = g(x) + j.$$

Since B is of period q, it leaves one element fixed. We may set $g(0) = 0$, and the above recurrence relation then takes the simple form $g(x) = xj$.

Then, corresponding to 2.2, we have the representation

$$2.5 \qquad A\colon\ x' = x + 1, \quad B\colon\ x' = jx, \qquad x = 0, 1, \ldots, p-1.$$

Further, letting g be that primitive root modulo p for which $g^2 = j$, B has the representation $x' = g^2x$ and, explicitly, our representations are

$$2.6 \qquad A = (0\,1\,2 \ldots p-1), \quad B = (1\,g^2\,g^4 \ldots g^{p-3})(g\,g^3 \ldots g^{p-2}).$$

The two cycles of B contain, respectively, quadratic residues and non-residues mod p.

It can be readily verified that the same group is generated no matter which primitive root g (that is, which number j) is considered.

Now we see from 1.5 that C stands in the same relation to B as does B to A. So B and C generate the metacyclic group $\{B, C\}$ of order qr. On q symbols, B and C can be represented as

$$2.7 \qquad B\colon\ y' = y + 1, \quad C\colon\ y' = m^s y, \qquad y = 0, 1, \ldots q-1,$$

or, explicitly,

$$2.8 \qquad \begin{aligned} B &= (0\,1\,2 \ldots q-1), \\ C &= (1\,m^s\,m^{2s} \ldots)(m\,m^{s+1} \ldots) \ldots (m^{s-1}\,m^{2s-1} \ldots), \end{aligned}$$

where $q = sr + 1$ and m is that primitive root mod q for which $m^s = l$.

We now see that the representations 2.5 and 2.7 can be combined. For in 2.5, B has the representation

$$(1\,g^2\,g^4 \ldots g^{p-3})(g\,g^3 \ldots g^{p-2}).$$

Let α be an arbitrary even integer. Then the cycle

$$(g^\alpha\,g^{\alpha+2}\,g^{\alpha+4} \ldots g^{\alpha+p-3})$$

is the same as $(1\,g^2\,g^4 \ldots g^{p-3})$, so that, for every α, we may set up a correspondence between 0 and g^α, 1 and $g^{\alpha+2}, \ldots, q-1$ and $g^{p-3+\alpha}$; that is,

$$z \rightarrow g^{\alpha+2z}.$$

In other words, the representation for B in 2.7 can be mapped on the first cycle of the representation for B in 2.5.

Then, in the representation for C in 2.7, in which $y' = m^s y$, we have

$$\begin{tikzcd} y \arrow[r] \arrow[d, leftrightarrow] & m^s y \arrow[d, leftrightarrow] \\ x = g^{\alpha+2y} \arrow[r] & g^{\alpha+2m^s y} = (x\,g^{-\alpha})^{m^s}.g^{\alpha} = g^{\alpha(1-m^s)}x^{m^s}. \end{tikzcd}$$

Now, let β be an arbitrary odd integer, and proceed as above with the second cycle of $B = (1\,g^2 \ldots g^{p-3})(g\,g^3 \ldots g^{p-2})$. We may again set up, for every β, a correspondence

$$z \to g^{\beta+2z},$$

which maps the representation for B in 2.7 on the second cycle of the representation for B in 2.5. Then, under the permutation C, $y' = m^s y$, we have

$$x \to g^{\beta(1-m^s)}x^{m^s}.$$

Now, for the first cycle of B in 2.5, that consisting of residues mod p, we set $z \to g^{\alpha+2z}$ and see that under C the same elements map as follows:

$$x \to g^{\alpha(1-m^s)}x^{m^s}.$$

For the second cycle of B, that consisting of non-residues mod p, we set $z \to g^{\beta+2z}$ and find that under C the elements follow the mapping

$$x \to g^{\beta(1-m^s)}x^{m^s}.$$

In other words, we demand that residues map on residues, non-residues on non-residues, and obtain the following representation on p symbols for C:

2.9
$$x' = g^{\alpha(1-m^s)}x^{m^s}, \qquad \epsilon = +1,$$
$$x' = g^{\beta(1-m^s)}x^{m^s}, \qquad \epsilon = -1.$$

It is easily shown that the choice of l among those numbers belonging to r mod q is an arbitrary one, and further that the same permutation C is obtained, with possibly a different order in the cycles, from different choices of pairs α and β. We may fix α, and then there will be q choices for β. So we obtain q permutations C. Indeed we might fix α at zero and have

2.10 $$C\colon\ x' = \tfrac{1}{2}(1+\epsilon)\,x^l + \tfrac{1}{2}(1-\epsilon)\,g^{\beta(1-l)}x^l.$$

C consists of $2s$ cycles of length r, and so is an even permutation. The elements 0, 1, and g^{β} remain fixed under C. A and B are even permutations, so the group $\{A, B, C\}$ consists entirely of even permutations.

In the above development of the permutation C, we required that residues mod p map on residues, non-residues on non-residues. This is not necessary; we might have obtained a representation for C by mapping residues on non-residues. C then takes the form

2.11 $$C\colon\ x' = \tfrac{1}{2}(1+\epsilon)\,g^{\beta}x^l + \tfrac{1}{2}(1-\epsilon)\,g^{-\beta l}x^l.$$

Since C consists of $2s$ cycles of length r, along with the transposition $(1\ g^{\beta})$, it is an odd permutation. Summarizing, we have

LEMMA 2.1. *The groups $\{A, B, C\}$ fall into two classes: those groups consisting entirely of even permutations generated with*

$$C: \begin{cases} x' = x^l, & \epsilon = +1; \\ x' = g^{\beta(1-l)}x^l, & \epsilon = -1; \end{cases}$$

and those groups consisting of even and odd permutations generated with

$$C: \begin{cases} x' = g^{\beta}x^l, & \epsilon = +1; \\ x' = g^{-\beta l}x^l, & \epsilon = -1. \end{cases}$$

We refer to these two classes as case I and case II, respectively, of our problem.

3. Certain relations among the groups $\{A, B, C\}$. For fixed p, q, and r, a set of $\frac{1}{2}(p+1)$ groups is obtained depending on the choice of β used in the permutation C. These groups are not all distinct up to isomorphism. The following lemma applies to these sets of groups.

LEMMA 3.1. (i) *In case* I, *for fixed p, the groups $\{A, B, C\}$ with $r = (q-1)/ks$ are respectively subgroups of the groups with $r = (q-1)/s$.*

(ii) *In case* II, *for fixed p, if the period of C is an odd multiple of the period of C', the groups $\{A, B, C'\}$ are respectively subgroups of the groups $\{A, B, C\}$.*

(iii) *In case* II, *for fixed p, the groups with $r = 2^{\lambda}\pi$, where $\lambda > 1$ and π is a product of powers of odd primes, contain as subgroups the corresponding groups from case* I, $r = 2$.

Consider part (i). This says that for $r = (q-1)/ks$ we have a set of $\frac{1}{2}(p+1)$ groups, and for $r = (q-1)/s$ we have a second set of $\frac{1}{2}(p+1)$ groups. There is a one-to-one correspondence between the two sets such that each group of the first set is a subgroup of one of the groups of the second set. An expression can be found relating these two sets, but this is not necessary for our application of the lemma. Similar statements apply to parts (ii) and (iii).

Part (i) follows by taking the kth power of the permutation C of period $r = (q-1)/s$,

$$C^k: \begin{cases} x' = x^{m^{ks}}, & \epsilon = +1, \\ x' = g^{\beta(1-m^{ks})}x^{m^{ks}}, & \epsilon = -1, \end{cases}$$

and pointing out that this is the permutation C obtained for $r = (q-1)/ks$.

In case II, r must necessarily be even if C, of order r, is to be odd. C is given by

$$C: \begin{cases} x' = g^{\beta}x^l, & \epsilon = +1, \\ x' = g^{-\beta l}x^l, & \epsilon = -1, \end{cases}$$

and is of order $(q-1)/s$. Taking powers of C we obtain

$$C^{2k}: \begin{array}{ll} x' = x^{l^{2k}}, & \epsilon = +1, \\ x' = g^{\beta(1-l^{2k})}x^{l^{2k}}, & \epsilon = -1, \end{array}$$

and

$$C^{2k+1}: \begin{array}{ll} x' = g^{\beta}x^{l^{2k+1}}, & \epsilon = +1, \\ x' = g^{-\beta l^{2k+1}}x^{l^{2k+1}}, & \epsilon = -1. \end{array}$$

Obviously, only odd powers of C yield another odd permutation. Since C^{2k+1} is simply the permutation C obtained with $r = (q-1)/(2k+1)s$, we see that part (ii) of our lemma follows.

Finally, consider case II with $r = 2^{\lambda}\pi$, $\lambda > 1$, π a product of powers of odd primes. C is now given by

$$C: \begin{array}{ll} x' = g^{\beta}x^{l}, & \epsilon = +1, \\ x' = g^{-\beta l}x^{l}, & \epsilon = -1, \end{array}$$

implying

$$C^{\frac{1}{2}r}: \begin{array}{ll} x' = x^{l^{r/2}} & \epsilon = +1, \\ x' = g^{\beta(1-l^{r/2})}x^{l^{r/2}}, & \epsilon = -1. \end{array}$$

where $l^r \equiv 1 \bmod q$, and hence $l^{\frac{1}{2}r} \equiv q - 1 \bmod q$. Then

$$\beta(1 - l^{\frac{1}{2}r}) \equiv \beta(2 - q) \bmod q.$$

As β runs over the odd numbers from 1 to $p - 2$, $\beta(2 - q) \bmod p - 1$ does also. Hence the permutation $C^{\frac{1}{2}r}$ will be the same permutation C as in case I for $r = 2$, viz.,

$$C: \begin{array}{ll} x' = x^{q-1}, & \epsilon = +1, \\ x' = g^{\beta}x^{q-1}, & \epsilon = -1, \end{array}$$

and part (iii) of the lemma is demonstrated.

4. Certain properties of the groups $\{A, B, C\}$.

THEOREM 4.1. *In case* I, $\mathfrak{H} = \{A, B, C\}$ *is a simple group. In case* II, $\mathfrak{H}$ *contains a simple subgroup of index two.*

THEOREM 4.2. *The order of* $\{A, B, C\}$ *is* $h = \delta pq(1 + np)$, *where* $\delta = 1$ *in case* I *and* $\delta = 2$ *in case* II.

It is known **(1; 4)** that if a permutation group of prime degree p contains an element A of period p, such that the only elements commuting with A are its own powers, then the order of the group is $p(p-1)(1+np)/t$, where p is the period of A, t the number of conjugate classes in $\{A\}$, and $1 + np$ the number of Sylow subgroups of order p.

Now certainly in a group on p symbols, any cycle of length p commutes only with its own powers. Hence the order of $\{A, B, C\}$ has the above form.

LEMMA 4.1. *There are at most two conjugate sets in* $\{A\}$.

For $B^{-1}AB = A^j$ and hence $A \sim \ldots \sim A^{j^{s-1}} \sim A^{j^s} = A, A^g \sim A^{gj} \sim \ldots$.

LEMMA 4.2. *In case* I, *there are exactly two conjugate sets in* $\{A\}$.

If $X^{-1}AX = A^g$, then, using the same procedure as in §2, X can be shown to have the form

$$X = (0 \;\; u \;\; u + gu \quad u + gu + g^2u \ldots);$$

that is, X is a cycle of order $p - 1$. Such a cycle is odd and not possible in case I.

Thus in case I, $t = 2$, and $h = pq(1 + np)$.

LEMMA 4.3. *The orders of the first and second commutator groups* $\mathfrak{H}'$ *and* $\mathfrak{H}''$ *contain the factor* p.

$A^{-1}B^{-1}AB = A^{j-1}$, so $\mathfrak{H}'$ contains a power, and hence all powers, of A. Also $B^{-1}C^{-1}BC = B^{l-1}$ and so $\mathfrak{H}'$ contains B. Thus $\mathfrak{H}' > \{A, B\}$, and pq divides h'. Similarly $\mathfrak{H}'' > \{A\}$, and the order of the second commutator subgroup, h'', is divisible by p.

LEMMA 4.4. *In case* I, *the commutator subgroup* $\mathfrak{H}'$ *is equal to* $\mathfrak{H}$.

According to Brauer we have three possibilities:

(1) $\mathfrak{H}$ has a normal subgroup of order $1 + np$; $t = p - 1$. This is impossible, since in case I, $t = 2$ and in case II, $t < 2$.

(2) $\mathfrak{H}$ has a normal subgroup of order $1 + np$; $t < p - 1$ and $\mathfrak{H}' \neq \mathfrak{H}''$, $\mathfrak{H}''$ has order $1 + np$, $\mathfrak{H}'$ has order $p(1 + np)$. This possibility is ruled out since, by the preceding lemma, $p|h''$.

We have then the third possibility holding, viz.,

(3) $\mathfrak{H}$ does not contain a normal subgroup of order $1 + np$, and $\mathfrak{H}'$ has order

$$h' = p\,\frac{p-1}{t'}\,(1 + np),$$

where $t|t', t'|p - 1$, and $t < t' < p - 1$. Further, $\mathfrak{H}' = \mathfrak{H}''$, and the group $\mathfrak{H}/\mathfrak{H}'$ is cyclic. Here t' denotes the number of classes of conjugate elements in H' which contain elements of order p.

Now in case I, $t = 2$; so t' is even, and, since $p - 1 = 2q$, with $t' < p - 1$, we have $t' = 2$. Thus $h' = h$ and $\mathfrak{H} = \mathfrak{H}'$.

LEMMA 4.5. *In case* II, *there is only one conjugate set in* $\{A\}$.

In case II, the commutator subgroup is a normal subgroup, proper or improper, of $\mathfrak{H}$. But it consists entirely of even permutations, and so $\mathfrak{H}'$ is properly contained in $\mathfrak{H}$. Now $t \neq 2$, since otherwise $\mathfrak{H} = \mathfrak{H}'$. So in case II, $t = 1$, and $h = 2pq(1 + np)$, completing the proof of Theorem 4.2.

LEMMA 4.6. *In case* II *the commutator subgroup* $\mathfrak{H}'$ *consists of all the even permutations in* $\mathfrak{H}$ *and has index two in* $\mathfrak{H}$.

For, the commutator subgroup has order

$$h' = p\frac{p-1}{t'}(1+np).$$

We proved $pq|h'$. Also $t|t'|p-1$, $t \leqslant t' < p-1$. The only possibilities for t' are 1, 2, q. But $t' \neq 1$, as $h' < h$. And $t' \neq q$, since $B^{-1}AB = A^s$ in $\mathfrak{H}'$ and so $t' \leqslant 2$. Thus $t' = 2$ and $\mathfrak{H}'$ is the subgroup of index 2 consisting of all even permutations.

Brauer has shown in the same paper that $\mathfrak{H}'$ is a simple group. But in case I, $\mathfrak{H}' = \mathfrak{H}$, and so $\mathfrak{H}$ is a simple group. In case II, $\mathfrak{H}''$ is a simple group and $\mathfrak{H}$ is like the symmetric group in that it contains a simple subgroup of index 2. Thus Theorem 4.1 is proved.

5. Case I, r even, $\beta = q$.

THEOREM 5.1. *The groups generated for* $r = 2$, $\beta = q$, *in case* I *are the alternating groups* $\mathfrak{A}_p$.

Here we have

$$C:\quad x' = \tfrac{1}{2}(g^\beta + 1)\, x^l - \tfrac{1}{2}(g^\beta - 1)\, x^{l+q},$$

and for $\beta = q$, $r = 2$, $l = q - 1$

$$C:\quad x' = x^{2q-1} = 1/x$$

$(x \neq 0)$. C leaves 0, 1 and -1 invariant.

Under the permutation $A^{-1}CA\; CA^{-1}C$, $x \to -x$ for $x \neq 0, 1$, and

$$0 \to -1, \quad -1 \to 1, \quad 1 \to 0.$$

This permutation, therefore, contains the cycle $(0\ {-1}\ 1)$, and the remaining elements form transpositions, since $A^{-1}CA\; CA^{-1}C$ is of period 2. Since $p = (p-3) + 3 = 2m + 3$, all elements are involved, and

$$A^{-1}CA\; CA^{-1}C = (0\ {-1}\ 1)(\text{product of transpositions}).$$

Squaring this permutation leaves the cycle $(0\ 1\ {-1})$. Applying Netto's Theorem **(5)**, "If a transitive group of degree n contains a circular permutation of prime order $q < \frac{2}{3}n$, then the group is either non-primitive, or it contains the alternating group," we see that $\mathfrak{H} = \mathfrak{A}_p$, since $\mathfrak{H}$ is primitive. (In this case, the presence of the cycle $(0\ 1\ 2 \ldots p-1)$ and the triplet $(0\ 1\ p-1)$ is sufficient for the proof that $\mathfrak{H} = \mathfrak{A}_p$.)

Now since the groups generated for even values of r with $\beta = q$ contain the groups generated for $r = 2$, $\beta = q$, from Lemma 3.1, part (i), we have at once

COROLLARY 5.1. *The groups generated for even values of* r, $\beta = q$, *case* I *are the alternating groups* $\mathfrak{A}_p$.

6. The groups $\{A, B, C\}$, case II. For fixed p, q is fixed, and we consider the possibilities for r. We have seen that in case II, r must be even. Then $r = 2$, an odd multiple of 2, or an even multiple of 2.

THEOREM 6.1. *If $r = 2$ in case* II, *the groups $\{A, B, C\}$ are the symmetric groups $\mathfrak{S}_p$.*

For in case II we may write

$$C: \quad x' = \tfrac{1}{2}(g^{\beta} + g^{-\beta l})\, x^{l} + \tfrac{1}{2}(g^{\beta} - g^{-\beta l})\, x^{l+q};$$

and for $r = 2$, $l = q - 1$, $g^{-\beta l} = g^{-\beta(q-1)} \equiv -\, g^{\beta}$; hence we have

$$C: \quad x' = g^{\beta} x^{2q-1}.$$

For $x \neq 0$, this can be written $x' = g^{\beta}/x$.
Consider the permutation

$$R = A^{-1}CA^{g^{\beta}}CA^{-1}: \quad x' = -\,1/x, \qquad\qquad x \neq 0, 1.$$

Under this permutation, $0 \to -\,1$, $-1 \to 1$, $1 \to 0$, and in standard form $R = (0\ {-1}\ 1) \cdot$ (product of transpositions). Then $R^2 = (0\ 1\ {-1})$, $A^{-1} R^2 A = (0\ 1\ 2)$; this permutation and $A = (0\ 1\ 2 \ldots p - 1)$ generate $\mathfrak{A}_p$. Since C itself is an odd permutation, $\{A, B, C\} = \mathfrak{S}_p$.

COROLLARY. *If r is an odd multiple of* 2 *in case* II *the groups $\{A, B, C\}$ are the symmetric groups $\mathfrak{S}_p$.*

This follows immediately from Lemma 3.1, part (ii), and Theorem 6.1.

THEOREM 6.2. *If $r = 2^{\lambda}\pi$, $\lambda > 1$, π odd, in case* II, *the groups $\{A, B, C\}$ are the symmetric groups $\mathfrak{S}_p$.*

It follows from Lemma 3.1, part (iii), that for fixed p, and for C with $r = 2^{\lambda}\pi$, the permutations $C^{\frac{1}{2}r}$ are the same as those obtained in case I with $r = 2$, viz.,

$$C^{\frac{1}{2}r}: \quad \begin{array}{ll} x' = x^{q-1}, & \qquad\qquad \epsilon = +\,1, \\ x' = g^{\beta}x^{q-1}, & \qquad\qquad \epsilon = -\,1. \end{array}$$

Then if $\beta = q$, $\{A, B, C\}$ must contain $\mathfrak{A}_p$, from Theorem 5.1; hence $\{A, B, C\}$ is $\mathfrak{S}_p$.

In the following we will assume $\beta \neq q$. In Lemma 4.5 we proved that in case II, there is only one conjugate set in $\{A\}$. Then $\mathfrak{H}$ must contain an element S such that $S^{-1} AS = A^{g}$. Such an S must be of the form

$$S: \quad x' = u + gx, \qquad\qquad u \text{ fixed},$$

and under $SA^{-u} B^{\frac{1}{2}(q-1)}$, $x \to -\, x$. That is, in case II, $\mathfrak{H}$ contains the permutation

$$T: x' = -\, x.$$

The permutation $C^{\frac{1}{2}r}$ leaves 0 fixed and, for $x \neq 0$, may be written

$$C^{\frac{1}{2}r}: \begin{array}{ll} x' = 1/x, & \epsilon = +1, \\ x' = -g^{\beta}/x, & \epsilon = -1. \end{array}$$

Then

$$C^{\frac{1}{2}r}T: \begin{array}{ll} x' = -1/x, & \epsilon = +1, \\ x' = g^{\beta}/x, & \epsilon = -1, \end{array}$$

maps quadratic residues mod p on quadratic non-residues, quadratic non-residues on quadratic residues, and leaves 0 fixed. Squaring this permutation we obtain

$$(C^{\frac{1}{2}r}T)^2: \begin{array}{ll} x' = g^{q+\beta}x, & \epsilon = +1, \\ x' = g^{q-\beta}x, & \epsilon = -1, \end{array}$$

$(x \neq 0)$. (Note here that if $\beta = q$, this permutation is the identity.) Finally,

$$(C^{\frac{1}{2}r}T)^2\, B^{\frac{1}{2}(q+\beta)}: \begin{array}{ll} x' = g^{2\beta}x, & \epsilon = +1, \\ x' = x, & \epsilon = -1, \end{array}$$

is a cyclic permutation of period q, viz.,

$$(1\, g^{2\beta}\, g^{4\beta} \ldots g^{2\beta(q-1)})$$

and applying Netto's theorem we obtain $\mathfrak{H} = \mathfrak{S}_p$.

7. Conclusion. It had been hoped that groups other than $\mathfrak{A}_p$, possibly multiply transitive groups such as the Mathieu groups $\mathfrak{M}_{11}$ and $\mathfrak{M}_{23}$, might appear in case I, but investigations have so far failed to produce any such groups.

References

1. R. Brauer, *On permutation groups of prime degree and related classes of groups*, Ann. Math., *44* (1943), 59–79.
2. R. D. Carmichael, *Abstract definitions of the symmetric and alternating groups and other permutation groups*, Quarterly J. Math., *49* (1923), 226–283.
3. ———, *Introduction to the theory of groups of finite order* (New York, Ginn and Co., 1937).
4. E. Mathieu, *Mémoire sur l'étude des fonctions de plusieurs quantités, sur la manière de les former et sur les substitutions qui les laissent invariables*, J. Math. pur. et app. (2), *6* (1861), 241–323.
5. E. Netto, *The theory of substitutions*, tr. F.N. Cole (Ann Arbor, Michigan, Inland Press, Register Publishing Co., 1892).
6. R. G. Stanton, *The Mathieu groups*, Can. J. Math., *3* (1951), 164–174.

Royal Military College of Canada

SET-TRANSITIVE PERMUTATION GROUPS

R. A. BEAUMONT AND R. P. PETERSON

1. Introduction. The concept of an s-ply transitive $(1 < s < n)$ permutation group on n symbols is of considerable importance in the classical theory of finite permutation groups, which was in the height of its development in the period around the turn of the century. The obvious generalization to a permutation group which is s set-transitive (i.e., a group which, for each pair of s-element unordered subsets S, T of the given n symbols, contains a permutation which carries S into T) seems to have received little attention. A discussion (**8**, p. 257) of the symmetry of an arbitrary n-person game leads in a natural way to the notion of a set-transitive permutation group (i.e., a group which is s set-transitive for all s) on the n players of the game. In the preface to (**8**), credit is given to C. Chevalley for solving the problem of determining all set-transitive groups. Since, to our knowledge, nothing has appeared in the literature on this problem, we believe that a complete and relatively simple solution is of interest.

In §2, the definitions are given, and the alternating and symmetric groups A_n and S_n, along with the trivial cases for $n < 3$, are considered. The properties of s set-transitive groups which are used for their enumeration are derived in §3, the principal results being contained in Theorems 5 and 6, which state that these groups are transitive and primitive. In this connection, a recent paper by Bays (**1**), relating a concept of primitivity for ordered pairs to the degree of transitivity of a permutation group, is of interest. In §4, a theorem on the distribution of prime numbers (**2**), is used to eliminate the possibility of set-transitive groups for $n > 82$. Various special results in (**5**) and (**7**) are used to obtain Theorem 10, which states that set-transitive groups, other than A_n and S_n, are possible only for $n = 5$, 6, and 9. Finally, in §5, all of the set-transitive groups of degree 5, 6, and 9 are determined. The results of this section are given in Theorem 11.

2. Definitions. We begin with a formal statement of the principal definitions and their immediate consequences. Since every permutation on n symbols sends the complete set of n symbols into itself, we exclude this trivial case from the following definition. Thus the case $n = 1$ is excluded. Further, the identity group I on n symbols can clearly be omitted from consideration, so that in the sequel, by a group $\mathfrak{G}$, we will mean a permutation group $\mathfrak{G} \neq I$ on the set N of $n > 2$ symbols, $N = [1, 2, \dots, n]$.

DEFINITION 1. A group $\mathfrak{G}$ is s *set-transitive* $(1 < s < n - 1)$ if for every pair of subsets S, T of $N = [1, 2, \dots, n]$, each containing s elements, there exists a permutation in $\mathfrak{G}$ which carries S into T.

Received October 9, 1952.

It should be noticed that according to this definition, "1 set-transitive" and "transitive" mean the same thing. We have the following immediate consequences of Definition 1:

(i) If the group $\mathfrak{G}$ contains an s set-transitive subgroup $\mathfrak{H}$, then $\mathfrak{G}$ is s set-transitive.

(ii) If the group $\mathfrak{G}$ is k-ply transitive, then $\mathfrak{G}$ is s set-transitive for all $s \leq k$.

(iii) If the group $\mathfrak{G}$ contains permutations which carry the set $S = [1,2, \ldots, s]$ into any other set T containing s elements, then $\mathfrak{G}$ is s set-transitive.

With $N = [1,2,3,4,5,6,7]$, the group $\mathfrak{G} = \{(1234567), (235)(476)\}$, is an example of a group which is 2 set-transitive but not doubly transitive. $\mathfrak{G}$ is not 3 or 4 set-transitive.

Definition 2. A group $\mathfrak{G}$ is *set-transitive* if $\mathfrak{G}$ is s set-transitive for all s $(1 \leq s \leq n-1)$.

Since the alternating group A_n $(n \geq 3)$ is $(n-2)$-ply transitive and since A_2 is intransitive, we have

Theorem 1. *The alternating group A_n is set-transitive except for $n = 2$. The symmetric group S_n is set-transitive.*

Proof. For $n \geq 3$, we have by (ii) that A_n is s set-transitive for all $s \leq n-2$. Since, in particular, A_n is transitive, A_n contains a permutation which sends n into any other symbol j. Therefore A_n contains a permutation which sends $S = [1, 2, \ldots, n-1]$ into any other $n-1$ element set T. By (iii), A_n is $n-1$ set-transitive, and hence set-transitive. Since S_n is n-ply transitive for all n, S_n is set-transitive.

It follows from the theorem that, in the determination of all set-transitive groups, we need only consider those groups $\mathfrak{G} \neq I$ which do not contain the alternating group. For $n = 2$, there are none. For $n = 3$, the only such groups are the three cyclic groups $\{(12)\}$, $\{(13)\}$, and $\{(23)\}$ which are not 1 or 2 set-transitive. Having disposed of these trivial cases, we will henceforth assume that $n \geq 4$.

3. Properties of s set-transitive groups. In this section we derive properties of s set-transitive groups which are needed for their enumeration. The principal result is that an s set-transitive group is primitive if $s > 1$.

Theorem 2. *A group $\mathfrak{G}$, which is the conjugate in S_n of an s set-transitive group $\mathfrak{H}$, is s set-transitive.*

Proof. Let σ be a permutation in S_n such that $\mathfrak{G} = \sigma \mathfrak{H} \sigma^{-1}$. Let $S = [1, 2, \ldots, s]$, and $J = [j_1, j_2, \ldots j_s]$ be an arbitrary s-element subset. Define the set $I = [i_1, i_2, \ldots, i_s]$ by $\sigma^{-1} S = I$, and the set $K = [k_1, k_2, \ldots, k_s]$ by $\sigma^{-1} J = K$. Since $\mathfrak{H}$ is s set-transitive, there exists a permutation $\tau \in \mathfrak{H}$ such that $\tau I = K$. Then

$$\sigma\tau\sigma^{-1} \in \mathfrak{G}, \quad \sigma\tau\sigma^{-1} S = \sigma\tau I = \sigma K = J.$$

THEOREM 3. *If the group $\mathfrak{G}$ is s set-transitive, then $\mathfrak{G}$ is $n - s$ set-transitive.*

Proof. Let $S = [1, 2, \ldots, n - s]$, $C(S) = [n - s + 1, n - s + 2, \ldots, n]$. Let $J = [j_1, j_2, \ldots, j_{n-s}]$ be an arbitrary $(n - s)$-element subset, and $C(J) = [i_1, i_2, \ldots, i_s]$ be the complement of J in N. There exists a permutation $\sigma \in \mathfrak{G}$ such that $\sigma C(S) = C(J)$, so that $\sigma(S) = J$.

THEOREM 4. *If the group $\mathfrak{G}$ is s set-transitive, then the order of $\mathfrak{G}$ is $m\binom{n}{s}$, where m is the order of the subgroup $\mathfrak{G}_1$ of $\mathfrak{G}$ consisting of those permutations which carry the subset $S = [1, 2, \ldots, s]$ into itself.*

Proof. It is clear that the subset $\mathfrak{G}_1$ of $\mathfrak{G}$ consisting of those permutations which carry $S = [1, 2, \ldots, s]$ into itself is a subgroup. Since there are $t = \binom{n}{s}$ distinct s-element subsets of N, denote them by $I_1 = S, I_2, \ldots, I_t$, and denote by $\sigma_1, \sigma_2, \ldots \sigma_t$ a set of permutations in $\mathfrak{G}$ such that $\sigma_k I_1 = I_k$, for $k = 1, 2, \ldots, t$. Then $\sigma_1, \sigma_2, \ldots, \sigma_t$ form a complete set of representatives of $\mathfrak{G}$ modulo $\mathfrak{G}_1$, so that the order of $\mathfrak{G}$ is mt, where m is the order of $\mathfrak{G}_1$, and $t = \binom{n}{s}$.

COROLLARY. *If the group $\mathfrak{G}$ is set-transitive, then the order of $\mathfrak{G}$ is divisible by the least common multiple of the binomial coefficients*

$$\binom{n}{1}, \binom{n}{2}, \ldots, \binom{n}{n-1}.$$

The 2 set-transitive group $\mathfrak{G} = \{(1234567), (235)(476)\}$ of degree 7 has order 21. Since $\binom{7}{2} = 21$, $\mathfrak{G}$ has minimum order for a 2 set-transitive group of this degree. Since

$$\binom{7}{3} = \binom{7}{4} = 35,$$

$\mathfrak{G}$ cannot be 3 or 4 set-transitive.

THEOREM 5. *If the group $\mathfrak{G}$ is s set-transitive for at least one s, then $\mathfrak{G}$ is transitive.*

Proof. Assume that $\mathfrak{G}$ is intransitive and let $L \subset N$ be the smallest transitivity set of $\mathfrak{G}$. Then L has l elements where $l \leqslant \frac{1}{2}n$, and we may assume by Theorem 3 that $l \leqslant s$. Since $l \leqslant s \leqslant n - 1$, there exists an s-element subset S of N such that $L \subseteq S$ and such that $C(S)$, the complement of S in N, contains an element not in L. By removing an element of L from S and replacing it by an element from $C(S)$, we can construct an s-element subset T such that $L \not\subseteq T$. Since $\mathfrak{G}$ is s set-transitive, there exists a permutation $\sigma \in \mathfrak{G}$ such that $\sigma S = T$. Since L is a transitivity set of $\mathfrak{G}$, $\sigma L = L$, and since $L \subseteq S$,

$\sigma L = L \subseteq \sigma S = T$. But this contradicts the choice of T, proving that $\mathfrak{G}$ must be transitive.

THEOREM 6. *If the group $\mathfrak{G}$ is s set-transitive for at least one $s > 1$, then $\mathfrak{G}$ is primitive.*

Proof. By Theorem 3, we may suppose that $s \leqslant \frac{1}{2}n$. $\mathfrak{G}$ is transitive by Theorem 5, and assume that $\mathfrak{G}$ is imprimitive. Then the set $N = [1, 2, \ldots, n]$ can be partitioned into $r \geqslant 2$ subsets N_i, each containing $l \geqslant 2$ elements, such that every permutation in $\mathfrak{G}$ carries each N_i into some N_j.

If $s \leqslant l$, there is an s-element subset S which is a subset of N_1. Since $s > 1$, there is an s-element subset T which contains elements from both N_1 and N_2. But since $\mathfrak{G}$ is s set-transitive, there exists a permutation in $\mathfrak{G}$ which carries S into T and this contradicts the assumption that $\mathfrak{G}$ is imprimitive.

If $l < s$, then since $s \leqslant \frac{1}{2}n$, we have $l < \frac{1}{2}n$ and $r \geqslant 3$. Then there is an s-element subset S which contains elements only from the sets of imprimitivity, $N_1, N_2, \ldots, N_k$ where $1 < k \leqslant \frac{1}{2}(r + 1)$, such that S contains $N_1, N_2, \ldots, N_{k-1}$. Since $r \geqslant 3$, $r - k \geqslant \frac{1}{2}(r - 1) \geqslant 1$, and the $r - k$ sets of imprimitivity, $N_{k+1}, \ldots, N_r$, are disjoint from those from which S was constructed. Therefore there is an s-element subset T, constructed by replacing one of the elements of N_1 by an element from N_{k+1}, which contains elements from $k + 1$ different sets of imprimitivity. Again, the existence of a permutation in $\mathfrak{G}$ which carries S into T contradicts the assumption that $\mathfrak{G}$ is imprimitive.

In the next section, we use a classical result (**3**, p. 199) due to Jordan and Netto, which connects primitivity with the degree of transitivity of $\mathfrak{G}$, to show that set-transitive groups are rare exceptions.

4. Determination of the values of n for which set-transitive groups may exist.

THEOREM 7. *If the group $\mathfrak{G}$ is s set-transitive $(s > 1)$, and if there exists a prime p such that $\frac{1}{2}n < p \leqslant n$ and such that p divides $\binom{n}{s}$, then $\mathfrak{G}$ is $(n-p+1)$-ply transitive.*

Proof. Since the order of $\mathfrak{G}$ is $m\binom{n}{s}$ by Theorem 4, if a prime p divides $\binom{n}{s}$, $\mathfrak{G}$ contains an element of order p. The elements of order p in $\mathfrak{G}$, when written as a product of cycles on disjoint letters, are products of cycles of length p. Since $p > \frac{1}{2}n$, the elements of order p in $\mathfrak{G}$ are cycles of length p. Such a cycle generates a cyclic subgroup $\mathfrak{H}$ of $\mathfrak{G}$ which is of degree p. Thus $\mathfrak{H}$ is primitive and leaves $n - p$ letters unchanged. Since $\mathfrak{G}$ is primitive by Theorem 6, we obtain the conclusion of the theorem by employing the result mentioned above (**3**, p. 199) which states that if a primitive group $\mathfrak{G}$ contains a primitive subgroup of degree m which leaves the remaining $n - m$ letters unchanged, then $\mathfrak{G}$ is $(n - m + 1)$-ply transitive.

COROLLARY 1. *If the group $\mathfrak{G}$ is s set-transitive $(s > 1)$, and if there exists a prime p such that* $\max(s, n - s) < p \leqslant n$, *then $\mathfrak{G}$ is $(n - p + 1)$-ply transitive.*

Proof. If $p > \max(s, n-s)$, then $p > \frac{1}{2}n$ and p divides

$$\binom{n}{s} = \frac{n!}{s!\,(n-s)!},$$

so that the hypotheses of the theorem are satisfied.

In the determination of set-transitive groups, the critical value of s is $s = [\frac{1}{2}n]$, where as usual this symbol denotes the greatest integer in $\frac{1}{2}n$.

COROLLARY 2. *If the group $\mathfrak{G}$ is $[\frac{1}{2}n]$ set-transitive, and if there exists a prime p such that $\frac{1}{2}(n+1) < p < n$, then $\mathfrak{G}$ is $(n-p+1)$-ply transitive.*

Proof. For $p > \frac{1}{2}(n+1) \geqslant [\frac{1}{2}(n+1)] = \max([\frac{1}{2}n], n - [\frac{1}{2}n])$, and the hypotheses of Corollary 1 are satisfied for $s = [\frac{1}{2}n]$.

We now make use of various known limits of transitivity to eliminate the possibility of the existence of $[\frac{1}{2}n]$ set-transitive, and therefore set-transitive, groups. The principal criterion is given in the following theorem.

THEOREM 8. *If there exists a prime p such that $\frac{1}{2}(n+1) < p < \frac{2}{3}n$, then a group $\mathfrak{G}$ on n symbols, which does not contain the alternating group A_n, cannot be $[\frac{1}{2}n]$ set-transitive.*

Proof. Assume that $\mathfrak{G}$ is $[\frac{1}{2}n]$ set-transitive. Then if a prime p exists in the given range, $\mathfrak{G}$ is $(n-p+1)$-ply transitive by Corollary 2, Theorem 7. But since $\frac{1}{3}n + 1 = n - \frac{2}{3}n + 1 < n - p + 1$, and since $\frac{1}{3}n + 1$ is an upper limit for the degree of transitivity (**3**, p. 152) for a group $\mathfrak{G}$ not containing A_n, we have a contradiction.

There are many refinements of Bertrand's postulate which states that a prime exists in the range between x and $2x$. One such result which is convenient for our purposes is due to Breusch (**2**). He shows that for $x \geqslant 48$, there always exists a prime between x and $9x/8$. For $n \geqslant 82$, $x = \frac{1}{2}(n+14) \geqslant 48$, and there exists a prime between $\frac{1}{2}(n+14)$ and $9(n+14)/16$. Since $\frac{1}{2}(n+1) < \frac{1}{2}(n+14)$, and $9(n+14)/16 < \frac{2}{3}n$ for $n \geqslant 82$, there exists a prime in the range given in the hypothesis of Theorem 8. By examining a table of primes, we find a prime strictly between $\frac{1}{2}(n+1)$ and $\frac{2}{3}n$ for all $n \geqslant 26$, and we have

THEOREM 9. *A group $\mathfrak{G}$ on $n \geqslant 26$ symbols, which does not contain the alternating group A_n, cannot be $[\frac{1}{2}n]$ set-transitive, and therefore is not set-transitive.*

For $n < 26$, a table of primes shows that a prime lies in the required range for $n = 8, 11, 12$ and for all n such that $17 \leqslant n \leqslant 24$. Since the cases for $n < 4$ have been previously discussed, we have only the cases

$$n = 4, 5, 6, 7, 9, 10, 13, 14, 15, 16, 25$$

to investigate. For these cases we first employ a better result on the limit of transitivity due to Miller (**7**, vol. III, p. 439) which states that if $n = kp + r$,

where p is a prime greater than the positive integer k and where $r > k$, then a group $\mathfrak{G}$ on n symbols, not containing the alternating group A_n, cannot be more than r-ply transitive, unless $k = 1$ and $r = 2$. As an example, with $n = 25$, we obtain from Corollary 2, Theorem 7 with $p = 17$, that if $\mathfrak{G}$ is 12 set-transitive, then $\mathfrak{G}$ is 9-ply transitive. But $25 = 1 \cdot 19 + 6$ with $k = 1$, $p = 19$, and $r = 6$, so that $\mathfrak{G}$ cannot be more than 6-ply transitive. Therefore there are no groups on $n = 25$ symbols, other than A_{25} and S_{25}, which are $[\frac{1}{2}n] = 12$ set-transitive. In this way, the cases $n = 10, 14, 15, 16$, and 25 are eliminated.

Miller has proved (**7**, vol. I, p. 200) that a transitive group of degree 13, which does not contain the alternating group A_{13}, is at most doubly transitive. By Corollary 2, Theorem 7 with $p = 11$, we have that if $\mathfrak{G}$ is $[\frac{1}{2}n] = 6$ set-transitive, then $\mathfrak{G}$ is triply transitive. Similarly, for $n = 7$, a transitive group not containing A_7 is at most doubly transitive (**5**, p. 186; **6**, p. 338; **7**, vol. I, pp. 1-9), while Corollary 2, Theorem 7, with $p = 5$, gives that if $\mathfrak{G}$ is $[\frac{1}{2}n] = 3$ set-transitive, then $\mathfrak{G}$ is triply transitive.

A 2 set-transitive group $\mathfrak{G}$ on $n = 4$ symbols has an order divisible by $\binom{4}{2} = 6$ by Theorem 4. The only such transitive groups are A_4 and S_4 (**6**, p. 338; **7**, vol. I, pp. 1-9). The cases $n = 2$ and 3 were eliminated in §1.

We summarize the above results in the following theorem.

THEOREM 10. *A group $\mathfrak{G}$ on n symbols, which does not contain the alternating group A_n, cannot be $[\frac{1}{2}n]$ set-transitive, and therefore is not set-transitive, with the exceptions of $n = 5, 6$, and 9.*

5. Determination of the set-transitive groups for $n = 5, 6$, and 9. In the following determination of set-transitive groups, we use the table of transitive groups on $n < 9$ symbols given by Cole in (**6**). Although this list has two omissions for $n = 8$, it has been verified by Miller (**7**, vol. I, pp. 1-9, 12-14) and others for $n = 5, 6$, and 9.

If a group $\mathfrak{G}$ on 5 symbols is 2 set-transitive, then by Theorem 4 the order of $\mathfrak{G}$ is divisible by $\binom{5}{2} = 10$. Thus the only possibilities for a 2 set-transitive group, other than A_5 or S_5, are the transitive groups:

$$G_1 = \{(12345), (1325)\}, \qquad \text{order } 20,$$
$$H_1 = \{(12345), (12)(35)\}, \qquad \text{order } 10,$$

and their conjugate subgroups in S_5. Since G_1 is doubly transitive, it is 2 set-transitive, and therefore 3 set-transitive by Theorem 3. Since G_1 is transitive, it is 4 set-transitive by this same theorem. Thus G_1 and its conjugate subgroups in S_5 (Theorem 2) are set-transitive. If H_1 were 2 set-transitive, then the order of H_1 would be $10m$, where m is the order of the subgroup of H_1 which carries the set [1,2] into itself. Thus $m = 1$, and the identity is the only

element of H_1 which carries [1,2] into itself. But $(12)(35) \in H_1$ carries [1,2] into itself. Therefore H_1 is not 2 set-transitive.

By Theorem 4, if a group $\mathfrak{G}$ on 6 symbols is 3 set-transitive, then the order of $\mathfrak{G}$ is $m\binom{6}{3} = 20m$, where m is the order of the subgroup of $\mathfrak{G}$ which sends the set [1,2,3] into itself. The only possibilities for a 3 set-transitive group, other than A_6 or S_6, are the transitive groups:

$$G_2 = \{(12345),(12)(35),(13465),(1325)\}, \qquad \text{order } 120,$$
$$H_2 = \{(12345),(12)(35),(13465)\}, \qquad \text{order } 60,$$

and their conjugate subgroups in S_6. Since G_2 is triply transitive, it is both 3 and 2 set-transitive, and by Theorem 3, G_2 is set-transitive, as are its conjugates in S_6. If H_2 were 3 set-transitive, then the order of the subgroup of H_2 which carries the set [1,2,3] into itself would be 3. However, the permutations

$$(1),(13)(45),(123)(465),(23)(56)$$

in H_2 carry [1,2,3] into itself. Therefore H_2 is not 3 set-transitive.

Again using Theorem 4, a 4 set-transitive group on 9 symbols has order $m\binom{9}{4} = 126m$. The only possibilities, other than A_9 and S_9, are the transitive groups:

$$G_3 = \{(1254673),\ (15)(29)(47)(68),\ (124)(765)\}, \qquad \text{order } 1512,$$
$$H_3 = \{(1254673),\ (15)(29)(47)(68)\}, \qquad \text{order } 504,$$

and their conjugate subgroups in S_9. Since H_3 is triply transitive, H_3 is 1, 2, 3, 6, 7 and 8 set-transitive by Theorem 3. By this same theorem if H_3 is 4 set-transitive, it is 5 set-transitive and consequently set-transitive. That a permutation can be found in H_3 which sends the set [1,2,3,4] into each of the 126 four element subsets of $N = [1,2,3,4,5,6,7,8,9]$ has been checked directly by the authors. Therefore H_3 is set-transitive, and since H_3 is a subgroup of G_3, G_3 is also set-transitive. We summarize these results in the following theorem which, as was indicated at the beginning of this section, depends in part on the correctness of the list of transitive groups (**6**, p. 338) for $n = 5$, 6, and 9.

THEOREM 11. *The only groups on n symbols, other than the symmetric and alternating groups S_n and A_n, which are $[\frac{1}{2}n]$ set-transitive, are the groups G_1, G_2, H_3, and G_3, and their conjugates, on* 5, 6, 9, *and* 9 *symbols, respectively. These four groups are set-transitive.*

In the verification that the group H_3 is 4 set-transitive, an element of the form $\tau\sigma^n$, where τ is an element of order two or three in H_3 and $\sigma = (1254673) \in H_3$, can be found which carries the set [1,2,3,4] into each four element subset $[a,b,c,d]$. For example, with $\tau = (64)(72)(51)(39)$ and $n = 4$, $\tau\sigma^n$ carries [1,2,3,4] into [2,3,5,9]. In the same way, with $\tau = (756)(412)(839)$ and $n = 0,1,4,5,6$, $\tau\sigma^n$ carries [1,2,3,4] into [1,2,4,9], [2,5,6,9], [1,6,7,9], [2,3,7,9], and [1,3,5,9], respectively.

It may be of some interest to give an additional description of the four set-transitive groups G_1, G_2, H_3, and G_3. As an abstract group, G_1 is metacyclic with defining relations $R^5 = S^4 = 1$, $S^{-1}RS = R^2$ (**7**, vol. III, p. 241). G_2 is isomorphic to the symmetric group S_5, and has the abstract defining relations $R^5 = S^4 = (RS^2)^3 = (R^2S)^2 = 1$ (**7**, vol. III, p. 241). The group H_3 is the simple group $LF(2,8)$ consisting of all the linear fractional transformations

$$x' = \frac{ax + b}{cx + d}$$

where a, b, c, d are elements of $GF(2^3)$ such that $ad - bc \neq 0$. As an abstract group, $H_3 = \{A, B\}$ has defining relations

$$A^7 = B^2 = (AB)^3 = (A^3BA^5BA^3B)^2 = 1$$

(**4**, p. 174). Finally, G_3 is isomorphic to the group of automorphisms of H_3, and $G_3 = \{A, B, C\}$, where C satisfies the relations

$$C^3 = 1, \; CAC^{-1} = A^2, \; CBC^{-1} = ABA^4BA^4BA.$$

Thus G_3 has order 1512 and contains H_3 as a normal subgroup of index 3. The generators A, B, and C are given by $A = (1254673)$, $B = (15)(29)(47)(68)$, and $C = (124)(765)$ in terms of the given generators of H_3 and G_3 as permutation groups on nine symbols.

References

1. S. Bays, *Sur la transitivité et la primitivité des groupes de substitutions*, Comm. Math. Helvet, *22* (1949), 17–30.
2. R. Breusch, *Zur Verallgemeinerung des Bertrandschen Postulates, das zwischen x und 2x stets Primzahlen liegen*, Math. Zeitschrift, *34* (1932), 505–526.
3. W. Burnside, *Theory of groups* (Cambridge, 1897).
4. ———, *Note on the simple group of order* 504, Math. Ann., *52* (1899), 174–176.
5. F. N. Cole, *Note on the substitution groups of six, seven, and eight letters*, Bull. New York Math. Soc., *2* (1893), 184–190.
6. ———, *List of the substitution groups of nine letters*, Quart. Jour. Math., *26* (1893), 372–388.
7. G. A. Miller, *Collected works*, vols. I and III (Urbana, 1935 and 1946).
8. J. von Neumann and O. Morgenstern, *Theory of games and economic behavior* (Princeton, 1947).

University of Washington

ON THE HOMOLOGY THEORY OF ABELIAN GROUPS

SAMUEL EILENBERG AND SAUNDERS MACLANE

1. Introduction. In **(1)** we have introduced the notions of "construction" and "generic acyclicity" in order to determine a homology theory for any class of multiplicative systems defined by identities. Among these classes the most interesting one is the class of associative and commutative systems Π with a unit element (containing the class of abelian groups). For this class we have exhibited a "cubical" construction $Q(\Pi)$ and proved its generic acyclicity. We have also indicated the initial stages (in low dimensions) of a second construction $A(\Pi)$. In the present paper, this gap is completed; the construction $A(\Pi)$ is defined (§4) and its generic acyclicity is proved (§5). The complexes $A(\Pi)$ are closely related to the complexes $A(\Pi, n)$ used in the theory of the groups $H(\Pi, n)$ **(2; 3)** which were introduced with a view to topological applications. It is this relation with the groups $H(\Pi, n)$ which was the initial motivation for the introduction of the notion of generic acyclicity.

We have found it convenient to employ extensively the notion of tensor product. In this connection we found a theorem characterizing the natural homomorphisms of tensor products, which, besides its own merits, has the effect of eliminating many explicit calculations that would otherwise be needed in the proofs of §5. This theorem on tensor products is stated in §2 and proved in §3. These two sections form a self-contained unit, which can be read independently of the rest of the paper.

2. Natural homomorphisms of tensor products. An *augmented commutative ring* R is a commutative ring together with a ring homomorphism α: $R \to J$ into the ring J of integers; an (augmented) ring homomorphism β: $R \to R'$ of two such rings is an ordinary ring homomorphism satisfying $\alpha'\beta = \alpha$. We recall here the convention that each ring has an identity element and that each ring homomorphism maps the identity into the identity.

Important examples of augmented rings are obtained by considering the integral ring $R = J(\Pi)$ of an associative and commutative system Π with a unit element. This ring $J(\Pi)$ is defined as the free abelian group generated by the elements $x \in \Pi$, with multiplication defined by the multiplication in Π and with augmentation given by $\alpha x = 1$. In particular, if Π is the free associative and commutative system with a set $\{y_\alpha\}$ as base, then $J(\Pi)$ is the free commutative ring with $\{y_\alpha\}$ as base, or equivalently $J(\Pi)$ is the polynomial ring $J[\{y_\alpha\}]$. The augmentation α assigns to each polynomial P the sum of its coefficients; in other words, $\alpha(P) = P(1)$.

Received April 3, 1954. This investigation was supported in part under contracts AF-18(600)-562 and Nonr-218(00).

If R_1 and R_2 are augmented commutative rings, their tensor product $R = R_1 \otimes R_2$ (over the integers) has the multiplication $(x_1 \otimes x_2)(x_1' \otimes x_2') = x_1x_1' \otimes x_2x_2'$ and the augmentation $\alpha(x_1 \otimes x_2) = (\alpha_1x_1)(\alpha_2x_2)$. If $\beta_1: R_1 \to R_1'$ and $\beta_2: R_2 \to R_2'$ are (augmented) ring homomorphisms then

$$\beta_1 \otimes \beta_2: R_1 \otimes R_2 \to R_1' \otimes R_2'$$

again is an (augmented) ring homomorphism. With this definition, $R_1 \otimes R_2$ constitutes a covariant functor of two variables.

We shall treat the tensor product as an associative operation and will consider the functors of t variables

$$R_1^{m_1} \otimes \dots \otimes R_t^{m_t} \tag{2.1}$$

where the kth factor is the m_k-fold tensor product $R_k \otimes \dots \otimes R_k$. We adopt the conventions that $R^0 = J$ and that $R \otimes J = R = J \otimes R$; this allows us freely to omit or add terms with exponent 0. In view of the natural isomorphism $R_1 \otimes R_2 \approx R_2 \otimes R_1$, the functors (2.1) are essentially the only ones that can be constructed using the tensor product.

We shall be concerned with the natural transformations

$$\rho: R_1^{m_1} \otimes \dots \otimes R_t^{m_t} \to R_1^{n_1} \otimes \dots \otimes R_t^{n_t} \tag{2.2}$$

of one of the functors (2.1) into another. Such a ρ is a family of maps, one for each t-tuple $(R_1, \dots, R_t)$, satisfying the following naturality condition: if $\beta_k: R_k \to Q_k (k = 1, \dots, t)$ are augmented ring homomorphisms then the diagram

$$\begin{array}{ccc} R_1^{m_1} \otimes \dots \otimes R_t^{m_t} & \xrightarrow{\rho} & R_1^{n_1} \otimes \dots \otimes R_t^{n_t} \\ \beta_1^{m_1} \otimes \dots \otimes \beta_t^{m_t} \downarrow & & \downarrow \beta_1^{n_1} \otimes \dots \otimes \beta_t^{n_t} \\ Q_1^{m_1} \otimes \dots \otimes Q_t^{m_t} & \xrightarrow[\rho]{} & Q_1^{n_1} \otimes \dots \otimes Q_t^{n_t} \end{array}$$

is commutative. In considering the transformations ρ three points of view are possible:

(1) Each map (2.2) is a group homomorphism; then ρ is called a ***natural group homomorphism***. With fixed $(m_1, \dots, m_t)$ and $(n_1, \dots, n_t)$ the natural group homomorphisms form an abelian group.

(2) Each map (2.2) is a ring homomorphism; then ρ is called a ***natural ring homomorphism***.

(3) Each map (2.2) is an augmented ring homomorphism; then ρ is called a ***natural augmented ring homomorphism***. Important examples of natural augmented ring homomorphisms (for one variable R) are the following:

$$\rho: R^m \to R^{m+1}, \quad \rho(x_1 \otimes \dots \otimes x_m) = x_1 \otimes \dots \otimes x_m \otimes 1, \tag{2.3}$$

$$\rho: R^{m+1} \to R^m, \quad \rho(x_1 \otimes \dots \otimes x_{m+1}) = (\alpha x_1)\, x_2 \otimes \dots \otimes x_{m+1}, \tag{2.4}$$

$$\rho: R^{m+1} \to R^m, \quad \rho(x_1 \otimes \dots \otimes x_{m+1}) = (x_1x_2) \otimes x_3 \otimes \dots \otimes x_{m+1}, \tag{2.5}$$

$$\rho: R^m \to R^m, \quad \rho(x_1 \otimes \dots \otimes x_m) = x_{\pi 1} \otimes \dots \otimes x_{\pi m}, \tag{2.6}$$

the latter for any permutation π of the digits $1, \dots, m$.

THEOREM I. *For given* $(m_1, \ldots, m_t)$ *and* $(n_1, \ldots, n_t)$, *the natural group homomorphisms* (2.2) *form a free abelian group with the set of natural ring homomorphisms as base. A natural ring homomorphism* ρ *may written uniquely as*

$$\rho = \rho_1 \otimes \ldots \otimes \rho_t$$

where $\rho_k\colon R_k^{m_k} \to R_k^{n_k}$ *is a natural ring homomorphism* $(k = 1, \ldots, t)$. *Each such* ρ_k *can be obtained by composition of the natural augmented ring homomorphisms of the type* (2.3)–(2.6).

COROLLARY. *All natural ring homomorphisms are augmented natural ring homomorphisms.*

COROLLARY. *The only natural ring homomorphism*

$$\rho\colon R_1^{m_1} \otimes \ldots \otimes R_t^{m_t} \to J$$

is the augmentation map.

Theorem I states in effect that natural group homomorphisms (2.2) are exactly those given by the obviously possible formulas. The proof will be given in the next section.

In all our exposition we have considered the ring J of integers as the "ground-ring"; this could be replaced by an arbitrary commutative ring K. Then R will be a K-algebra, $\alpha\colon R \to K$ will be a K-algebra homomorphism and all tensor products will be tensor products over K. Our proof of Theorem I necessitates the assumption that K is an integral domain of characteristic zero.

3. Proof of Theorem I. The proof employs two elementary lemmas on polynomial identities.

LEMMA 1. *If* $x_1, \ldots, x_n$ *are independent indeterminates over an integral domain* D *of characteristic* 0, *and if the polynomial* $f \in D[x_1, \ldots, x_n]$ *satisfies the identity*

$$(3.1)\qquad \begin{aligned} &f(x_1 + y_1 - z_1, \ldots, x_n + y_n - z_n) \\ &\qquad = f(x_1, \ldots, x_n) + f(y_1, \ldots, y_n) - f(z_1, \ldots, z_n) \end{aligned}$$

then f *has degree at most* 1.

Proof. Set $g(x_1, \ldots, x_n) = f(x_1, \ldots, x_n) - f(0, \ldots, 0)$. Then g satisfies

$$g(x_1 + y_1, \ldots, x_n + y_n) = g(x_1, \ldots, x_n) + g(y_1, \ldots, y_n).$$

This is the usual identity characterizing a homogeneous linear polynomial.

LEMMA 2. *If a polynomial* $f \in D[x_1, \ldots, x_n]$ *with coefficients in an integral domain* D *satisfies the identity*

$$(3.2)\qquad f(x_1y_1, \ldots, x_ny_n) = f(x_1, \ldots, x_n)\, f(y_1, \ldots, y_n),$$

then f *is a monomial or zero.*

Here and in the sequel the word "monomial" will be understood as "monomial with coefficient 1."

Proof. The identity yields

$$f(x_1, \ldots, x_n) = f(x_1, 1, \ldots, 1) f(1, x_2, \ldots, x_n).$$

Since $f(x_1, 1, \ldots, 1)$ and $f(1, x_2, \ldots, x_n)$ each satisfy the identity (3.2) for 1 and $n - 1$ variables respectively, the proposition is reduced to the case $n = 1$. For this case, if $f \neq 0$, set $f(x) = ax^m + g(x)$, with $a \neq 0$ and $g(x)$ of degree k less than m. Then $f(xy) = f(x)f(y)$ becomes

$$ax^m y^m + g(xy) = a^2 x^m y^m + ax^m g(y) + ay^m g(x) + g(x)g(y).$$

Comparison of coefficients yields $a = a^2$ and hence $a = 1$, since D is an integral domain. Consideration of terms of degree $m + k$ then shows that $g(x) = 0$. Thus $f(x) = x^m$, as desired.

We first prove Theorem I for $t = 1$. As a "model" ring for this proof we use the polynomial ring

$$P = J[y_1, \ldots, y_m], \qquad \alpha(y_i) = 1,$$

in the independent indeterminates $y_1, \ldots, y_m$ with integer coefficients. Let

$$\kappa_j: P \to P^n, \qquad j = 1, \ldots, n,$$

be the natural injection of P into the jth factor of P^n. The ring P^n will be identified with the polynomial ring $J[Y]$ in the set

$$Y = \{\kappa_1 y_1, \ldots, \kappa_n y_1; \ldots; \kappa_1 y_m, \ldots, \kappa_n y_m\}$$

of mn independent indeterminates $\kappa_j y_i$, for $i = 1, \ldots, m$ and $j = 1, \ldots, n$, each with augmentation 1. A polynomial in the ring $J[Y]$ will be called *generic* if it is of degree at most one in each of the m strings of variables $\kappa_1 y_i, \ldots, \kappa_n y_i$.

With each natural group homomorphism $\rho: R^m \to R^n$ we associate a polynomial $f_\rho \in J[Y]$ as follows:

$$f_\rho = \rho(y_1 \otimes \ldots \otimes y_m), \quad \rho: P^m \to P^n.$$

THEOREM I′. *The correspondence $\rho \to f_\rho$ establishes an isomorphism between the group of natural group homomorphisms and the subgroup of generic polynomials in $J[Y]$. Furthermore ρ is a natural ring homomorphism if and only if f_ρ is a monomial.*

Theorem I′ includes the first part of Theorem I, for $t = 1$. The remaining two parts of Theorem I will also be established in the course of the proof. We begin with

LEMMA 3. *If $x_i \in R$ $(i = 1, \ldots, m)$ are elements with $\alpha(x_i) = 1$, then*

$$\rho(x_1 \otimes \ldots \otimes x_m) = f_\rho(\kappa_j x_i).$$

The last symbol is to be interpreted as the value of the polynomial f_ρ when each of the indeterminates $\kappa_j y_i$ is replaced by the element $\kappa_j x_i$ of R^n.

Proof. Let $\beta\colon P \to R$ be the (unique) augmented ring homomorphism satisfying $\beta(y_i) = x_i$ $(i = 1, \ldots, m)$. Since β^m is a ring homomorphism we obtain

$$\begin{aligned}\rho(x_1 \otimes \ldots \otimes x_m) &= \rho\,\beta^m(y_1 \otimes \ldots \otimes y_m) = \beta^n f_\rho(\kappa_j y_i)\\ &= f_\rho(\beta\kappa_j y_i) = f_\rho(\kappa_j \beta y_i) = f_\rho(\kappa_j x_i).\end{aligned}$$

We are now ready to tackle the proof of Theorem I′. It is clear that the correspondence $\rho \to f_\rho$ is a homomorphism. To show that it is a monomorphism assume $f_\rho = 0$. Then, by Lemma 3, $\rho(x_1 \otimes \ldots \otimes x_m) = 0$ whenever the elements x_i have augmentation 1. Now for any augmented ring R and any element $x \in R$ we have

$$x = (\alpha x - 1)\,\theta + (x + (1 - \alpha x)\,\theta),$$

where θ denotes the identity element of R. Since θ and $x + (1 - \alpha x)\,\theta$ both have augmentation 1, it follows from the multilinearity of the tensor product that $\rho = 0$.

Next we show that the polynomial f_ρ is generic. To simplify the notation, we shall show only that f_ρ is of degree at most 1 in the string of variables $\kappa_1 y_1, \ldots, \kappa_n y_1$. Let Z be the set of the remaining indeterminates of Y and let $D = J[Z]$. Then D is an integral domain of characteristic zero and we may identify $J[Y]$ with $D[\kappa_1 y_1, \ldots, \kappa_n y_1]$. Consider the ring

$$Q = P[u, v] = J[y_1, \ldots, y_m, u, v],$$

where the additional indeterminates u and v also have augmentation 1. Let $w = y_1 + u - v$; then $\alpha w = 1$. We apply Lemma 3 to the ring Q. Regarding f_ρ as a polynomial with coefficients in D, we obtain

$$\rho(w \otimes y_2 \otimes \ldots \otimes y_m) = f_\rho(\kappa_1 w, \ldots, \kappa_n w).$$

If on the left-hand side we replace w by $y_1 + u - v$, expand by linearity, and again apply Lemma 3, we obtain

$$f_\rho(\kappa_1 y_1, \ldots, \kappa_n y_1) + f_\rho(\kappa_1 u, \ldots, \kappa_n u) - f_\rho(\kappa_1 v, \ldots, \kappa_n v).$$

Thus $f_\rho \in D[\kappa_1 y_1, \ldots, \kappa_n y_1]$ satisfies the identity (3.1) of Lemma 1, and therefore f_ρ is of degree at most 1 in the string of variables $\kappa_1 y_1, \ldots, \kappa_n y_1$.

Next, suppose that ρ is a natural ring homomorphism. We apply ρ to the ring $R = J[u_1, \ldots, u_m, v_1, \ldots, v_m]$, where $\alpha u_i = \alpha v_i = 1$. The definition of the product in R^m yields

$$(u_1 \otimes \ldots \otimes u_m)(v_1 \otimes \ldots \otimes v_m) = u_1 v_1 \otimes \ldots \otimes u_m v_m.$$

Then applying Lemma 3 and the fact that ρ is a ring homomorphism we obtain

$$f_\rho((\kappa_j u_i)(\kappa_j v_i)) = f_\rho(\kappa_j u_i)\, f_\rho(\kappa_j v_i).$$

Since $\kappa_j u_i$ and $\kappa_j v_i$ are arbitrary indeterminates, Lemma 2 implies that f_ρ is a monomial.

To complete the proof of Theorem I′ we must show that each generic monomial $f \in J[Y]$ is of the form f_ρ for some ρ. Observe that those natural ring homomorphisms $\rho: R^m \to R^n$ which are composites of homomorphisms of the type (2.3)–(2.6) may be described as follows. Let $A_0 \cup \ldots \cup A_n$ be a decomposition of the set $\{1, \ldots, m\}$ into disjoint sets A_j. Then

$$\rho(x_1 \otimes \ldots \otimes x_m) = \alpha(x_{A_0})\, x_{A_1} \otimes \ldots \otimes x_{A_n},$$

where x_{A_j} is the product of the x_i's with $i \in A_j$; and $x_{A_j} = 1$ if $A_j = 0$. The monomial f_ρ of this homomorphism contains the indeterminate $\kappa_j y_i$ if and only if $x_i \in A_j$. Clearly any generic monomial may be obtained in this way from a suitable decomposition $A_0 \cup \ldots \cup A_n$.

The proof of Theorem I for general t is like that above for $t = 1$, with the use of t "model" rings

$$P_k = J[y_{k,1}, \ldots, y_{k,m_k}], \qquad \alpha(y_{k,i}) = 1; \; k = 1, \ldots, t.$$

For any ρ as in (2.2) we again have f_ρ in $J[Y]$, where Y is now a set of $\sum m_k n_k$ independent indeterminates $\kappa_j y_{k,i}$; $k = 1, \ldots, t, j = 1, \ldots, n_k, i = 1, \ldots, m_k$, and

$$f_\rho = \rho(y_{1,1} \otimes \ldots \otimes y_{1,m_1} \otimes \ldots \otimes y_{t,1} \otimes \ldots \otimes y_{t,m_t}).$$

In concluding the proof, we observe that each generic monomial f in the indeterminates Y admits a unique factorization $f = f_1 \ldots f_t$, where each f_k is a generic monomial in the indeterminates $\kappa_j y_{k,i}$ with k fixed. Each f_k determines a natural ring homomorphism $\rho_k: R_k^{m_k} \to R_k^{n_k}$ and $\rho = \rho_1 \otimes \ldots \otimes \rho_t$ satisfies $f_\rho = f_{\rho_1} \ldots f_{\rho_t} = f_1 \ldots f_t = f$.

4. The complexes $A^n(R)$. Let R be an augmented commutative ring. We convert R into a graded ring by assigning the degree zero to all the elements of R. Further by introducing a differentiation (i.e., boundary operator) which is identically zero we convert R into an augmented graded ∂-ring. We may now apply the bar construction **(2**, §7**)** and define

$$A(R, 1) = B(R).$$

From now on we continue by applying the normalized bar construction **(2**, §12**)** and define

$$A(R, n) = B_N(A(R, n - 1)), \qquad n > 1.$$

The elements of $A(R, n)$ may be described as follows. For each sequence of integers $k_1, \ldots, k_{r-1}, 1 < k_i < n$, consider the r-fold tensor product $T_r(R) = R \otimes \ldots \otimes R$. A typical element of this tensor product will be written as

$$(4.1) \qquad [x_1|_{k_1} x_2|_{k_2} \ldots |_{k_{r-1}} x_r]$$

instead of the usual $x_1 \otimes x_2 \otimes \ldots \otimes x_r$. The degree (i.e., dimension) of the element (4.1) is $n + k_1 + \ldots + k_{r-1}$. As a graded group $A(R, n)$ is the direct sum of all these tensor products for all sequences $k_1, \ldots, k_{r-1}$ as above and of the group J whose elements have degree zero. Thus $A(R, n)$ has no elements

of degrees $0 < d < n$ while, in degree n, $A(R, n)$ consists of R itself. In addition, $A(R, n)$ is equipped with a boundary operator ∂ and a product $*_n$ which are natural and which convert $A(R, n)$ into a graded ∂-ring. The complexes $A(R, n)$ for various n are compared by means of the suspension map

$$S: A(R, n) \to A(R, n+1)$$

which is a monomorphism of the additive group structure, raises the dimension by 1, and anticommutes with the boundary operator.

The details of the definition of $A(R, n)$ were motivated by geometrical applications. Indeed, if Π is an abelian group, then the homology groups of $A(J(\Pi), n)$ are those of a space X with vanishing homotopy groups $\pi_i(X)$ except for $\pi_n(X) = \Pi$. For algebraic purposes it is convenient to introduce the complexes $A^n(R)$ obtained from $A(R, n)$ by the following three modifications: (i) The group J in degree zero is removed. (ii) All degrees are lowered by $n - 1$; the element (4.1) has thus degree $1 + k_1 + \ldots + k_{r-1}$. (iii) The boundary operator ∂ is replaced by $\partial' = (-1)^n\partial$.

With these changes the suspension operator $S: A^n(R) \to A^{n+1}(R)$ preserves degree and boundary and may be regarded as an inclusion. We thus obtain the nested sequence of complexes

$$A^1(R) \subset A^2(R) \subset \ldots \subset A^n(R) \subset \ldots,$$

whose union will be denoted by $A(R)$. In this complex we have no product; each of the subcomplexes $A^n(R)$ has a product $*_n$ (inherited from $A(R, n)$) which, however, does not have a unit element.

The complexes $A^n(R)$ and $A(R)$ commence in degree 1 with R while in degree 2 we have the group $R \otimes R$ with boundary given by

$$\partial[x_1|_1x_2] = (\alpha x_2)[x_1] - [x_1x_2] + (\alpha x_1)[x_2]. \tag{4.2}$$

The 1-dimensional homology group thus is the quotient of R by the subgroup generated by the elements (4.2). We may denote this quotient by $h(R)$ and adjoin it to the complex $A^n(R)$ (or $A(R)$) in dimension zero, with the natural map $R \to h(R)$ as boundary operator. The resulting complex will be called the augmented complex $A^n(R)$ (or $A(R)$); its homology groups are trivial in dimensions <2.

If Π is a commutative associative system with a unit element and $R = J(\Pi)$ then it will be customary to write $A(\Pi, n)$, $A^n(\Pi)$, $A(\Pi)$, and $h(\Pi)$ instead of $A(J(\Pi), n)$, etc.

In particular, let Π be a free system with base $\{y_\alpha\}$. Then $R = J(\Pi)$ is the ring of polynomials $J[y_\alpha]$. Let $\sigma = [x_1|_{k_1} \ldots |_{k_{r-1}}x_r]$ be an element of $A(\Pi, n)$ (or of $A^n(\Pi)$ or of $A(\Pi)$) with $x_1, \ldots, x_r \in \Pi$. We shall say that σ is *generic* if the product $x_1 \ldots x_r$ is a monomial of degree $\leqslant 1$ in each variable y_α; i.e., if no generator y_α is repeated in $x_1 \ldots x_r$. The unit element of $A(\Pi, n)$ also is regarded as generic. The generic elements span a subgroup which will be denoted by $A(y_\alpha; n)$ (or $A^n(y_\alpha)$ or $A(y_\alpha)$). The complexes considered here

are direct sums of tensor products and the boundary operator is a natural group homomorphism, so Theorem I may be applied. It asserts that the boundary operator can be obtained from linear combinations of tensor products of composites of the special homomorphisms (2.3)–(2.6). The formulas displayed for these special homomorphisms clearly show that they carry generic elements into generic elements, hence the same applies to the boundary operator. In other words, the subgroup $A^n(y_\alpha, n)$ (or $A^n(y_\alpha)$ or $A(y_\alpha)$) is stable under the boundary operator, and thus forms a subcomplex called the *generic* subcomplex of $A(\Pi, n)$ (or of $A^n(\Pi)$ or $A(\Pi)$, respectively). The generic complexes $A^n(y_\alpha)$ and $A(y_\alpha)$ may be augmented by adjoining in dimension zero the generic subgroup $h(y_\alpha)$ of $h(\Pi)$ which is defined as the image of the generic subgroup of $J(\Pi)$ under the natural homomorphism $J(\Pi) \to h(\Pi)$.

The determination of the homology groups of the generic subcomplexes is our main objective here. The results are contained in the three theorems below. In stating these results it is convenient to assume that the base $\{y_\alpha\}$ of the free system Π is indexed by a simply ordered set $\{\alpha\}$.

THEOREM II. *The homology groups of the generic complex $A(y_\alpha; n)$ are as follows: H_0 is the infinite cyclic group generated by the unit element of $A(\Pi, n)$: $H_{nr}(r = 1, 2, \ldots)$ is the free abelian group generated by the homology classes of the cycles*

$$[y_{\alpha_1}]*_n \ldots *_n[y_{\alpha_r}], \qquad \alpha_1 < \ldots < \alpha_r.$$

In other dimensions the homology groups are zero.

THEOREM IIa. *The homology groups of the generic complex $A^n(y_\alpha)$ are as follows: $H_{n(r-1)+1}$ $(r = 1, 2, \ldots)$ is the free abelian group generated by the homology classes of the cycles*

$$[y_{\alpha_1}]*_n \ldots *_n[y_{\alpha_r}], \qquad \alpha_1 < \ldots < \alpha_r.$$

In other dimensions the homology groups are zero. In the augmented generic complex the group H_1 (corresponding to $r = 1$ above) is zero while the remaining groups are unchanged.

THEOREM IIb. *The homology groups of the generic complex $A(y_\alpha)$ are zero, except in dimension* 1, *where the group H_1 is the free abelian group generated by the homology classes of the cycles $[y_\alpha]$. In the augmented generic complex all the homology groups are zero.*

This last result implies that $A(\Pi)$ is a generically acyclic (augmented) construction in the sense of **(1)**. Hence the homology of $A(\Pi)$ is naturally isomorphic to that of the generically acyclic cubical construction on Π, as described in **(1)**.

In the complex $A(R)$, $[x_1|_{n+1}x_2]$ is a chain of the lowest dimension not contained in $A^n(R)$. Since this chain has dimension $n + 2$, it follows that

$$H_q(A(R)) = H_q(A^n(R)) = H_{q-1+n}(A(R, n)), \qquad q < n + 1.$$

This implies that for an abelian group Π (or more generally, for an associative and commutative system Π with a unit element) the homology groups $H_q(\Pi) = H_q(A(\Pi))$, furnished by the theory of generic acyclicity, coincide with the groups $H_{q-1+n}(\Pi, n) = H_{q-1+n}(A(R, n))$ $(q < n + 1)$. The latter are the *stable* groups (under suspension) in the theory of the groups $H(\Pi, n)$ corresponding to the "jump" $q - 1$.

5. Proofs of results of §4. We first observe that in Theorems IIa and IIb the second halves, dealing with the augmented complexes, are immediate consequences of the first halves. It is also clear that Theorem II implies Theorem IIa.

Next we derive Theorem IIb from Theorem IIa. Since $A(y_\alpha)$ is the union of the increasing sequence of complexes $A^n(y_\alpha)$ the homology group $H_q(A(y_\alpha))$ is the limit of the direct sequence of groups

$$H_q(A^1(y_\alpha)) \rightarrow H_q(A^2(y_\alpha)) \rightarrow \ldots \rightarrow H_q(A^n(y_\alpha)) \rightarrow \ldots$$

where the maps are induced by inclusion (i.e., suspension). For $q = 1$, it follows from Theorem II that the maps are isomorphisms and thus give the desired description of $H_1(A(y_\alpha))$. If $q > 1$ then for n sufficiently large, q is not of the form $n(r - 1) + 1$. Thus by Theorem IIa, $H_q(A^n(y_\alpha)) = 0$ for all n sufficiently large, and therefore $H_q(A(y_\alpha)) = 0$.

Thus Theorem II is the only one that still requires a proof. Since all the operations involved commute with direct limits, it is clear that it suffices to prove Theorem II in the case the (simply ordered) set $\{y_\alpha\}$ is a finite sequence $\{y_1, \ldots, y_t\}$.

Let $E(y_1, \ldots, y_t; n)$ be the free abelian group generated by the symbols

(5.1) $$y_{i_1}y_{i_2}\ldots y_{i_r}, \quad 1 \leq i_1 < i_2 < \ldots < i_r \leq t, \quad 0 \leq r.$$

The symbol corresponding to $r = 0$ will be denoted by 1. We introduce a grading in $E(y_1, \ldots, y_t; n)$ by assigning to the element (5.1) the degree rn. We also introduce a boundary operator which is identically zero. We shall prove that

(5.2) *The correspondence*

$$\phi(y_{i_1} \ldots y_{i_r}) = [y_{i_1}]*_n \ldots *_n[y_{i_r}]$$

is a chain equivalence

$$\phi = E(y_1, \ldots, y_t; n) \rightarrow A(y_1, \ldots, y_t; n).$$

This will clearly imply that the homology groups of $A(y_1, \ldots, y_t; n)$ are as described in Theorem II.

We shall now apply the "tensor product theorem" of **(3)** to reduce the proof of (5.2) to the case $t = 1$. The "tensor product theorem" establishes a chain equivalence

$$A(R_1, n) \otimes A(R_2, n) \underset{f}{\overset{g}{\rightleftarrows}} A(R_1 \otimes R_2, n),$$

with natural g and f and natural homotopies Φ: $gf \simeq$ ident., Ψ: $fg \simeq$ ident. Further g has the form $g(\sigma \otimes \tau) = \sigma *_n \tau$, provided that we agree to regard $A(R_1, n)$ and $A(R_2, n)$ as subcomplexes of $A(R_1 \otimes R_2, n)$ in the obvious way. Since f, g, Φ, and Ψ are natural, they must have the form prescribed in Theorem I. Take $R_i = J(\Pi_i)$, where Π_1 is the free system with base $y_1, \ldots, y_{t-1}$ and Π_2 is that with base y_t. These maps will then carry generic elements to generic elements, and therefore we have an induced chain equivalence

$$A(y_1, \ldots, y_{t-1}; n) \otimes A(y_t; n) \underset{f}{\overset{g}{\rightleftarrows}} A(y_1, \ldots, y_t; n).$$

for $t > 1$. Now consider the commutative diagram

$$\begin{array}{ccc} E(y_1, \ldots, y_{t-1}; n) \otimes E(y_t; n) & \xrightarrow{h} & E(y_1, \ldots, y_t; n) \\ \downarrow \phi_1 \otimes \phi_2 & & \downarrow \phi \\ A(y_1, \ldots, y_{t-1}; n) \otimes A(y_t; n) & \xrightarrow[g]{} & A(y_1, \ldots, y_t; n), \end{array}$$

where

$$h(y_{t_1} \cdots y_{t_r} \otimes 1) = y_{t_1} \cdots y_{t_r}$$

and

$$h(y_{t_1} \cdots y_{t_r} \otimes y_t) = y_{t_1} \cdots y_{t_r} y_t,$$

and where ϕ_1 and ϕ_2 are the maps of (5.2). Clearly h is an isomorphism. If therefore ϕ_1 and ϕ_2 are chain equivalences then so is $g(\phi_1 \otimes \phi_2) = \phi h$. Thus ϕ also is a chain equivalence.

The above argument reduces the proof to the case $t = 1$. To establish Theorem II in this case we employ the "Normalization theorem" of **(2)**. An element

$$[x_1|_{k_1} \cdots |_{k_{r-1}} x_r]$$

of $A(R, n)$ is called a *norm* if $r > 1$ and $x_i = 1$ for at least one $i = 1, \ldots, r$. The norms generate a subcomplex; factoring $A(R, n)$ by this subcomplex we obtain the *normalized* complex $A_N(R, n)$. In **(2**, Theorems 4.1 and 12.1**)** we have established a chain equivalence

$$A(R, n) \underset{g}{\overset{f}{\rightleftarrows}} A_N(R, n)$$

where $fg =$ identity, f is the natural factorization map and both g and the homotopy Φ: $gf \simeq$ ident. are natural. Furthermore $gf \equiv$ ident., modulo norms.

Passing to the generic complexes $A(y_\alpha; n)$, we introduce the normalized complex $A_N(y_\alpha; n)$ just as above and denote by f': $A(y_\alpha; n) \to A_N(y_\alpha; n)$ the natural factorization homomorphism. The map gf: $A(R, n) \to A(R, n)$ and the homotopy Φ: $gf \simeq$ ident. are both natural, and hence by Theorem I induce a map k: $A(y_\alpha; n) \to A(y_\alpha; n)$ and a homotopy Φ': $k \simeq$ ident. Since $\ker f' \subset \ker k$, it follows that k admits a factorization $k = g'f'$ with g': $A_N(y_\alpha; n) \to A(y_\alpha; n)$.

Since gf and hence k are congruent to the identity, modulo norms, it follows that $f'g'$ = ident. We therefore obtain a chain equivalence

$$A(y_a;\, n) \underset{g'}{\overset{f'}{\rightleftarrows}} A_N(y_a;\, n).$$

Now consider the generic complex $A(y;\, n)$ for a single generator y. In this complex all the chains of dimension $>n$ are norms. Since $\partial[y] = 0$ it follows that the complexes $A_N(y;\, n)$ and $E(y;\, n)$ may be identified. Under this identification we have $f'\phi$ = ident. Therefore

$$g' = g'f'\phi \simeq \phi,$$

so that ϕ also is a chain equivalence. This completes the proof.

References

1. Samuel Eilenberg and Saunders MacLane, *Homology theories for multiplicative systems*, Trans. Amer. Math. Soc., *71* (1951), 294–330.
2. ———, *On the groups $H(\Pi, n)$*, I, Ann. Math., *58* (1953), 55–106.
3. ———, *On the groups $H(\Pi, n)$*, II, Ann. Math., *60* (1954), 49–139.

Columbia University *The University of Chicago*

A CANONICAL SET FOR MATRICES OVER A PRINCIPAL IDEAL RING MODULO m

L. E. FULLER

1. Introduction. If $m \in P$ where P is a p.i.r. (principal ideal ring), then $P/\{m\}$ is a commutative ring with unit element. The elements of this ring are designated by $\bar{a}$ where $a \in P$. The set of square matrices of order n with elements in $P/\{m\}$ forms a ring with unit element. The units in this ring are the unimodular matrices, i.e., the matrices whose determinants are units of $P/\{m\}$. By the following definition, these unimodular matrices determine equivalence classes in the ring of matrices.

Two matrices A and B are *row equivalent*, or *left associates*, if there exists a unimodular matrix U such that $UA = B$.

We shall derive a canonical set for our matrices under row equivalence. This was first done by the author for P the ring of integers **(1)**. The present paper simplifies and extends that result.

2. The canonical set. The basic case to be considered is for $m = p^k$, p a prime in P, and k a positive integer. An element $a \in P$ and p^k have a g.c.d. of the form p^t where $0 \leqslant t \leqslant k$. We designate $d(a) = t$ as the *degree* of the element a. All elements of the same degree are associates in $P/\{p^k\}$. If an element is of degree zero, it has an inverse in $P/\{p^k\}$ and hence is a unit. If $d(a) = k$, then the element is in the zero class in the modular ring. All other elements are proper divisors of zero in $P/\{p^k\}$. The elements $\bar{a}$ of $P/\{p^k\}$ thus belong to one of $k + 1$ ordered classes. The order of a class is determined by the degree of any element in it. For convenience, the bar over the elements of $P/\{p^k\}$ will be dropped when they are elements of a matrix.

THEOREM 1. *Every nth order matrix with elements in $P/\{p^k\}$ is the left associate of a matrix having the following properties*:

1. The degree of every element is at least equal to the degree of the diagonal element of its row, that is, $d(a_{rs}) \geqslant d(a_{rr})$ for all r and s.

2. The degree of every element above the diagonal is greater than the degree of the diagonal element of its row unless that diagonal element is zero, i.e., $d(a_{rs}) > d(a_{rr})$ for all $s > r$ if $a_{rr} \neq 0$.

3. Every diagonal element is of the form p^t, $0 \leqslant t \leqslant k$.

4. If for $r \neq s$, $d(a_{rr}) \geqslant d(a_{ss})$, then $a_{rs} = 0$. If $d(a_{ss}) > d(a_{rr})$, then $a_{rs} \in P/\{p^k\}/\{a_{ss}\}$, that is, a_{rs} is unique modulo a_{ss}.

Properties 1 and 2 give the relationship between the elements in a row and their diagonal element. Property 4 does this for the elements in a column and

Received July 27, 1953; in revised form May 31, 1954.

their diagonal element. Property 3 uses a convenient representative for the classes of elements of the same degree.

The second property could be applied to the elements below the diagonal instead of to those above. For uniqueness, it is necessary to have it one way or the other. For the exception noted in Property 2, $a_{rs} = 0$ for all s by Property 1. This would be true for either form of 2.

If $k = 1$, then $P/\{p^k\}$ is a field. The elements are either of degree 1 (if $= 0$) or of degree 0. By Property 3 all diagonal elements are either zero or unity. If the diagonal element is zero, the rest of the row is zero by 1. If the diagonal element is of degree 0, the rest of the column is zero by 4. By Property 2, the matrix is triangular with zeros above the diagonal. These properties are those of the Hermite form for a square matrix over a field.

Before proving the theorem, let us outline briefly how the diagonal elements are to be chosen. Each element of the matrix belongs to one of the $k + 1$ ordered classes based on their degrees. We consider those in the class of lowest degree for the given matrix. For the first diagonal element, we choose any one in this class with one exception. When a row contains more than one element of the class of lowest degree, our choice in property 2 requires that we consider only the element with the highest column index. In step 2 we form a submatrix by deleting the column used in step 1 and the corresponding row. From this subarray, a choice of a second element in made in exactly the same manner. At each succeeding step, one more column and the corresponding row are deleted. An element is then selected from the resulting square submatrix as in the first step.

Step 1. By an interchange of rows, if necessary, place the chosen element in the diagonal position. This element can then be changed to a power of p by using a suitable unit as a multiplier. Because of the way in which the diagonal element was selected, all other elements of the column are multiples of it. By elementary transformations these can be reduced to zero. The index of this column and the corresponding row will be designated as the $\bar{1}$st.

Step 2. Make a choice of a new element from the submatrix formed without the $\bar{1}$st row and column. Place this element in the diagonal position and change it to a power of p. The elements in this column will, as before, be multiples of this diagonal element so that they can be reduced to zero. The one possible exception to this is the element in the $\bar{1}$st row. It can be transformed to the representative of its residue class modulo this new diagonal element. The index of this column and the corresponding row will be designated as the $\bar{2}$nd.

In general, the $\bar{h}$th diagonal element is chosen from the submatrix formed by the deletion of the rows and columns designated as $\bar{j}$th for $j = 1, \ldots, h - 1$. The selected element is placed in the diagonal and transformed to a power of p. All elements in rows not designated as yet are multiples of this diagonal element and can be reduced to zero. The elements in designated rows belong to residue classes modulo this $\bar{h}$th diagonal element. Each can then be transformed to the representative of its class.

Working with elements of least degree ensures Property 1. The choice of the element with the higher column index gives us Property 2. Properties 3 and 4 obviously follow from what is done after the selection of a diagonal element.

3. Uniqueness of the canonical set. To prove uniqueness, we assume that two canonical matrices A and B are in the same row equivalence class. Under this assumption, there exists a unimodular matrix Q such that $QA = B$. We must prove that $A = B$. This is done by an induction on the columns of A and B in the matrix equation. For the induction, we shall order the columns of A first according to the degrees of their diagonal elements, starting with the columns whose diagonal elements are of least degree. Then for those of the same degree, we order according to their indices, starting with the largest one. The degrees of the diagonal elements thus form a non-decreasing sequence under the specified ordering. There is always an increase in the degree whenever the column index is increased. This ordering is similar to that used in choosing the diagonal elements, the only difference being in the final ordering of the indices. This is done to simplify the proof. Because of this similarity, the bar notation will again be used to designate the ordering. At each step the degrees of the elements involved in the equations will play a key role. The *reductio ad absurdum* proof will come from an equation with the left side having all terms of higher degree than the single term on the right side. This situation will arise as a consequence of certain combinations of Property 2 and the following lemma, where it is assumed $a_{\bar{h}\bar{h}} \neq 0$.

LEMMA 1. *If for* $i \geqslant h$:

$$\begin{aligned} &\text{(a)}\ \bar{h} \geqslant \bar{\imath}, \quad \textit{then } d(a_{\bar{\imath}\bar{\imath}}) \geqslant d(a_{\bar{h}\bar{h}}), \\ &\text{(b)}\ \bar{h} < \bar{\imath}, \quad \textit{then } d(a_{\bar{\imath}\bar{\imath}}) > d(a_{\bar{h}\bar{h}}). \end{aligned}$$

This is the symbolic statement of the ordering on the columns of A that we are using.

LEMMA 2. *If* $i > h$, *then* $a_{\bar{\imath}\bar{h}} = 0$.

By Property 1 and Lemma 1, the following inequalities hold:

$$d(a_{\bar{\imath}\bar{h}}) \geqslant d(a_{\bar{\imath}\bar{\imath}}) \geqslant d(a_{\bar{h}\bar{h}}).$$

The conclusion is a consequence of Property 4.

Since Q is unimodular, it must have at least one unit in every row and column. We shall see that the only elements of degree zero will be the identity elements in the diagonal unless some diagonal element of A is zero. Then the corresponding column of Q will be arbitrary.

THEOREM 2. *The canonical set is unique.*

The $\bar{1}$st column of A contains all zeros except possibly the diagonal element, by Lemma 2. If it were also zero, then all diagonal elements of A would be zero by Lemma 1. Then by Property 1, A would be the zero matrix. It follows

that B must also be the zero matrix. In this event we have $A = B$ at once. In case the diagonal element is non-zero, the equations for this column take the simple form:

$$q_{\bar r \bar 1}\, a_{\bar 1 \bar 1} = b_{\bar r \bar 1}, \qquad r = 1, \ldots, n.$$

Since Q is unimodular, some $q_{\bar r \bar 1}$ must be of degree zero. If $d(q_{\bar 1 \bar 1}) = 0$, then $d(a_{\bar 1 \bar 1}) = d(b_{\bar 1 \bar 1})$. By Property 3, $a_{\bar 1 \bar 1} = b_{\bar 1 \bar 1}$ so that $q_{\bar 1 \bar 1}$ is the identity element. Consequently, $d(b_{\bar r \bar 1}) \geqslant d(b_{\bar 1 \bar 1})$ for all r. Therefore by Property 4, $b_{\bar r \bar 1} = 0$ for $r \neq 1$. This means that $q_{\bar r \bar 1}\, a_{\bar 1 \bar 1} = \delta_{\bar r \bar 1}\, a_{\bar 1 \bar 1}$, so that only the $q_{\bar 1 \bar 1}$ is a unit in the $\bar 1$st column of Q. Because $d(a_{\bar 1 \bar s}) \geqslant d(a_{\bar 1 \bar 1})$, $q_{\bar r \bar 1}\, a_{\bar 1 \bar s} = \delta_{\bar r \bar 1}\, a_{\bar 1 \bar s}$ for all s by Property 1.

If $q_{\bar 1 \bar 1}$ were not a unit, then some other $q_{\bar r \bar 1}$ would be of degree zero. We can derive a contradiction to this assumption by considering the equation involving the $\bar r$th diagonal of B:

$$\sum_{\bar i} q_{\bar r \bar i} a_{\bar i \bar r} = b_{\bar r \bar r}.$$

We know that

$$d(a_{\bar i \bar r}) \geqslant d(a_{\bar i \bar 1}) \geqslant d(a_{\bar 1 \bar 1}) = d(b_{\bar r \bar 1}) \geqslant d(b_{\bar r \bar r}).$$

The first and last inequalities hold by Property 1. The strict equality is a result of the assumption on $q_{\bar r \bar 1}$. The other inequality follows by Lemma 1. We shall now see that at least one of the three $\geqslant$'s is a strict inequality for each $\bar i$. This gives a false equation with all terms on the left of higher degree than the one on the right.

If $\bar 1 > \bar r$, the last inequality is a strict inequality by property 2 of the canonical set. If $\bar r > \bar 1$ and $\bar i > \bar 1$, the second $\geqslant$ is now strict by Lemma 1(b), since $i > 1$. If $\bar r > \bar 1$ and $\bar 1 \geqslant \bar i$, then $\bar r > \bar i$ so that the first $\geqslant$ cannot be an equality, by Property 2.

Assume that the $\bar j$th column of A is the same as the corresponding column of B for $j = 1, \ldots, h - 1$. We shall now show this to be true for the $\bar h$th column of A. From our assumption we know that for all $j < h$,

$$q_{\bar r \bar j}\, a_{\bar j \bar s} = \delta_{\bar r \bar j}\, a_{\bar j \bar s}.$$

By Lemma 2, we also known that $a_{\bar r \bar h} = 0$ for $r > h$. These assumptions mean that the equations involving the elements in the $\bar h$th column take on one of two forms:

$$q_{\bar r \bar h}\, a_{\bar h \bar h} = b_{\bar r \bar h}, \qquad \text{if } r > h,$$
$$q_{\bar r \bar h}\, a_{\bar h \bar h} + a_{\bar r \bar h} = b_{\bar r \bar h}, \qquad \text{if } r < h.$$

In case $a_{\bar h \bar h} = 0$, then $a_{\bar r \bar h} = b_{\bar r \bar h}$ for all $r < h$. This is also true for $r > h$, since $a_{\bar r \bar h} = 0 = b_{\bar r \bar h}$, so that the $\bar h$th columns are the same. By Property 1, $a_{\bar h \bar s} = 0$ for all s so that $q_{\bar r \bar j}\, a_{\bar j \bar s} = \delta_{\bar r \bar j}\, a_{\bar j \bar s}$ will hold for all $j \leqslant h$. In addition, $a_{\bar r \bar r} = 0$ for all $r > h$, by Lemma 1.

If $a_{\bar h \bar h} \neq 0$, then some $q_{\bar r \bar h}$ in the first equation must be a unit, since $q_{\bar j \bar j}$ is the identity element for $j < h$. If it is $q_{\bar h \bar h}$, then, as before, $a_{\bar h \bar h} = b_{\bar h \bar h}$ by Property 3 and $q_{\bar h \bar h} = 1$. In the second equation $a_{\bar h \bar h}$ can be replaced by $b_{\bar h \bar h}$ so that, by Property 4, $a_{\bar r \bar h} = b_{\bar r \bar h}$ for all r.

If $q_{\bar{h}\bar{h}}$ were not of degree zero, then some other $q_{\bar{r}\bar{h}}$, $r > h$, must be a unit. We again obtain a contradiction to this assumption from the equation involving the $\bar{r}$th diagonal element of B. In this equation, if $i < h$, then

$$q_{\bar{r}\bar{i}}\, a_{\bar{i}\bar{r}} = \delta_{\bar{r}\bar{i}}\, a_{\bar{i}\bar{r}} = 0$$

by the induction hypothesis. This reduces the equation to the form:

$$\sum_{i > h} q_{\bar{r}\bar{i}} a_{\bar{i}\bar{r}} = b_{\bar{r}\bar{r}}.$$

We know that

$$d(a_{\bar{i}\bar{r}}) \geqslant d(a_{\bar{i}\bar{i}}) \geqslant d(a_{\bar{h}\bar{h}}) = d(b_{\bar{r}\bar{h}}) \geqslant d(b_{\bar{r}\bar{r}}).$$

In exactly the same manner as before, one of the three inequalities is found to be a true inequality. The only difference lies in using h instead of 1. Therefore this is a false equation and the two columns are identical.

4. Extension of the results. We first note that the Chinese Remainder Theorem (**2**, p. 18) holds for a p.i.r.

LEMMA 3. *If m_1 and m_2 are relatively prime elements of a p.i.r., and the $\bar{a}_i$ are any given residue classes modulo m_i, then there exists a residue class $\bar{a}$ modulo $m_1 m_2$ such that $a \equiv a_i$ modulo m_i, for $i = 1,2$.*

Since m_1 and m_2 are relatively prime, in the p.i.r. there exists an f and g such that

$$fm_1 + gm_2 = 1.$$

This means that f is the inverse of m_1 modulo m_2, and similarly that g is the inverse of m_2 modulo m_1. We can now define our desired residue class using the following representative:

$$a = a_1 gm_2 + a_2 fm_1.$$

That the residue class thus determined has the required properties can be readily verified.

With one more lemma we can outline the extension to the case of the modulus $m = p_1^{k_1} \cdot \ldots \cdot p_n^{k_n}$, where the p_i are distinct primes. These prime powers $p_i^{k_i}$ will be called *primary factors* of m. The canonical set we choose will not be easy to describe but it will be unique because the residue class determined in Lemma 3 is unique.

LEMMA 4. *If A and B are row equivalent over $P/\{m\}$, they are row equivalent over $P/\{p^k\}$, where p^k divides m.*

If A and B are row equivalent, then there exists a matrix U such that $UA = B$. Since the determinant of U is a unit in $P/\{m\}$, it must be prime to m. If it is prime to m, it is prime to p^k, and hence is a unit in $P/\{p^k\}$.

The converse of this lemma is not necessarily true, since a number can be prime to p^k and not to m. However, any number that is prime to all primary factors of m is prime to m.

The lemma tells us that all row equivalent matrices over $P/\{m\}$ are row equivalent over $P/\{p^k\}$ for each primary factor of m. For a given class over $P/\{m\}$, there will correspond a row equivalence class over $P/\{p^k\}$. For each of the primary factors of m, the corresponding equivalence class has a canonical representative. One can find by Lemma 3 a matrix over $P/\{m\}$ that is congruent to each of these representatives. This matrix will be row equivalent to the given class over $P/\{m\}$, since it is row equivalent for every divisor of m. We call it the canonical matrix of the given equivalence class.

References

1. L. E. Fuller, *The Hermite canonical form for a matrix with elements in the ring of integers modulo m*, Thesis, University of Wisconsin, 1950.
2. C. C. MacDuffee, *Introduction to abstract algebra* (New York, 1940).

THE WEDDERBURN THEOREM

HARRY GOHEEN

Wedderburn, in 1905, proved that there are no finite skew-fields **(5)**. Wedderburn's result has also been proved by Dickson, Artin, Witt, and Zassenhaus **(2**; **1**; **6**; **7)**; however, it seems to the author that the proofs so far given introduce concepts not obviously related to the theorem. It is the purpose of this note to use a result of Cartan, which was later proved in greater generality by Hua **(4)**, to give a simpler and more direct version of the proof of Zassenhaus.

Lemma 1 (Hua). *Any invariant division subring of a skew field must be contained in the center of the skew field.*

The proof, being well known, is omitted.

Let $N(H)$ denote the normalizer of a subgroup H of a finite group G. We have

Lemma 2 (Zassenhaus). *If all the elements of $N(H_i)$ commute with all the elements of H_i for all Abelian subgroups H_i of G, then G is Abelian.*

Proof.[1] Let the lemma be true for any finite group K in which the hypotheses hold, provided K has fewer elements than the finite group M. Suppose that the hypotheses hold for M. We distinguish two possible cases: (1) M has a non-trivial center Z and (2) the center of M is the identity. In case (1) the quotient group M/Z has fewer elements than M and the hypotheses hold for M/Z. Hence M/Z is Abelian, and if x and y are any two elements in M,

$$xy = yxz$$

for some element z in Z. Then y is in the normalizer of the Abelian group generated by x and Z. By hypothesis, y commutes with every element of this Abelian subgroup, and $yx = xy$. Since x and y are arbitrary, M is Abelian.

In case (2), let C_1 be a maximal subgroup in M of index $n > 1$. Then C_1 is Abelian. If C_1 is invariant, every element of M may be written as t^ic_j, for t an element of M not in C_1 and c_j an element of C_1. Then, if u and v are two elements of M,

$$uv = (t^ic_j)(t^kc_l).$$

Because of the commutativity of c_j and c_l, of t^i and t^k, and of t and any element of C_1, this may be written as $(t^kc_l)(t^ic_j)$. Hence $uv = vu$.

Received February 25, 1954.

[1]This is a modification of the proof given in **(7)**.

On the other hand, if in case (2) C_1 is its own normalizer, then C_1 has n conjugates in M, and M has a faithful representation as a permutation group. The permutation corresponding to x in M is

$$\sigma(x) = \begin{pmatrix} C_1 & C_2 & \ldots & C_n \\ x^{-1}C_1x & x^{-1}C_2x & \ldots & x^{-1}C_nx \end{pmatrix}.$$

In this representation, the subgroup fixing the letter C_1 is the subgroup C_1. The representation is of class $n - 1$; for if a permutation $\sigma(z)$ fixes two letters C_i and C_j, z must be in C_i and C_j. Then since both C_i and C_j are Abelian and together generate M, z is in the center of M and therefore is the identity. Since the representation is of class $n - 1$, then by a theorem of Frobenius on such groups **(3)**, M has an invariant subgroup A consisting of the elements corresponding to permutations of degree n in the representation. Furthermore, A is Abelian, and A and C_1 generate M. Thus every element of M may be written in the form a_ic_j with a_i from A and c_j from C_1. Then, if u and v are any two elements of M,

$$uv = (a_ic_j)(a_kc_l) = vu,$$

since A and C_1 are both Abelian and every element of C_1 is commutative with every element of A. Since u and v are arbitrary elements, M must be Abelian. Thus case (2) cannot occur, for we have the contradiction that Z is M itself. To complete the induction we note that the lemma is true for all groups of prime order.

See p 413

THEOREM. *There is no finite skew field.*

Proof. Let D be a finite skew field, M its multiplicative group, H any Abelian subgroup of M, and Z the center of D. By an induction argument, it will be demonstrated that every element from the normalizer of H commutes with every element of H. By Lemma 2, it will follow that M is Abelian.

If H is in Z, its normalizer is M and, since H is in the center, every element of M commutes with every element of H. Let H be an Abelian subgroup of M which is not in Z. Then H and Z generate a commutative subfield X of D. It cannot be D itself since D is a skew field. However X is contained in a proper maximal division subring C of D.

Now we make the induction assumption that all finite division rings which have fewer elements than D must be fields. Then C is a field and its multiplicative group is in the normalizer of H in M. However, no element not in C can be in the normalizer of H; for if s is, then X is invariant under s and hence under the division ring generated by C and s. But this division ring is D and X is therefore an invariant division subring of D. Then, by Lemma 1, X is in Z. This contradicts the assumption that H is not in Z. Therefore, the multiplicative group of C is the normalizer of H in M. Since C is a field, every element of the normalizer of H in M commutes with every element of H. By Lemma 2, M is Abelian and D is a field, contrary to hypothesis. The hypothesis that D

is a skew field being contradictory, it follows that there are no finite skew fields with as many elements as D if there are none with fewer elements.

To complete the induction it suffices to note that any division ring with a prime number of elements must be a field.

References

1. E. Artin, *Ueber einen Satz von Herrn J. H. Maclagan-Wedderburn*, Abh. Math. Sem. Univ. Hamburg, *5* (1927), 245–250.
2. L. E. Dickson, *On finite algebras*, Gött. Nachr. (1905), 358–393.
3. G. Frobenius, *Ueber auflösbare Gruppen*, IV, Berl. Ber., (1901), 1223–1225.
4. L. K. Hua, *Some properties of a sfield*, Proc. Nat. Acad. Sci. U.S.A., *35* (1949), 533–537.
5. J. H. Maclagan-Wedderburn, *A theorem on finite algebras*, Trans. Amer. Math. Soc., *6* (1905), 349–352.
6. E. Witt, *Ueber die Kommutativität endlicher Schiefkörper*, Abh. Math. Sem. Univ. Hamburg, *8* (1930), 413.
7. H. J. Zassenhaus, *A group-theoretic proof of a theorem of Maclagan-Wedderburn*, Proc. Glasgow Math. Assoc., *1* (1952), 53–63.

Iowa State College

THE EXISTENCE OF A DISTRIBUTION FUNCTION FOR AN ERROR TERM RELATED TO THE EULER FUNCTION

PAUL ERDÖS AND H. N. SHAPIRO

1. Introduction. The average order of the Euler function $\phi(n)$, the number of integers less than n which are relatively prime to n, raises many difficult and still unanswered questions. Thus, for

$$R(x) = \sum_{n \leq x} \phi(n) - \frac{3}{\pi^2} x^2, \tag{1.1}$$

and

$$H(x) = \sum_{n \leq x} \frac{\phi(n)}{n} - \frac{6}{\pi^2} x, \tag{1.2}$$

it is known that $R(x) = O(x \log x)$ and $H(x) = O(\log x)$. However, though these results are quite old, they were not improved until recently. Walfisz **(1)** has given the outline of a proof of

$$R(x) = O(x(\log x)^{3/4}(\log \log x)^2).$$

On the other hand it is known **(3)** that

$$R(x) \neq O(x \log \log \log x). \tag{1.3}$$

and

$$H(x) \neq O(\log \log \log x). \tag{1.4}$$

In this direction it was proved in **(4)** that each of the following inequalities holds for infinitely many integral x (c a certain positive constant):

$$R(x) > cx \log \log \log \log x, \tag{1.5}$$

$$R(x) < - cx \log \log \log \log x, \tag{1.6}$$

$$H(x) > c \log \log \log \log x, \tag{1.7}$$

$$H(x) < - c \log \log \log \log x. \tag{1.8}$$

In this paper we propose to continue the study of the error function $H(x)$, and will prove that $H(x)$ possesses a continuous distribution function. By this we mean that for $N(n, u)$ = the number of $m < n$ such that $H(m) > u$, we have for each u, $-\infty < u < \infty$, that the limit

$$\lim_{n \to \infty} \frac{N(n, u)}{n} = F(u) \tag{1.9}$$

exists; and the non-increasing function $F(u)$ is continuous for all u.

In the case of additive arithmetic functions, necessary and sufficient conditions for the existence of a distribution function are known **(5; 6)**. The methods used in **(5)** to establish the sufficient conditions seem to apply in a fairly general way for establishing the existence of a continuous distribu-

Received May 19, 1954.

tion function even for a function which is not additive (7). This method serves also as the basic framework of the proof given here for the existence of a continuous distribution function for $H(x)$.

There are essentially three steps. *First*, we introduce for each integer $k > 1$, the function

$$H_k(x) = \sum_{n<x} \frac{\phi((n, A_k))}{(n, A_k)} - x\prod_{p<p_k}\left(1 - \frac{1}{p^2}\right), \tag{1.10}$$

where

$$A_k = \prod_{p<p_k} p;$$

where p_k is the kth prime. It is then shown that for each u, with fixed k, if $N_k(n, u)$ is the number of $m < n$ such that $H_k(x) > u$, the limit

$$\lim_{n\to\infty} \frac{N_k(n, u)}{n} = F_k(u) \tag{1.11}$$

exists. We then see that (1.9) follows if we can show that, for a given u and any $\epsilon > 0$, the inequality

$$|N(n, u) - N_k(n, u)| < \epsilon n \tag{1.12}$$

holds for each $k > k_0 = k_0(\epsilon)$ and all $n > n_0 = n_0(k)$. For from (1.12) we have

$$\left|\sup \frac{N(n, u)}{n} - \lim_{n\to\infty} \frac{N_k(n, u)}{n}\right| < \epsilon, \qquad \left|\inf \frac{N(n, u)}{n} - \lim_{n\to\infty} \frac{N_k(n, u)}{n}\right| < \epsilon,$$

for $k > k_0$. This in turn gives

$$\left|\sup \frac{N(n, u)}{n} - \inf \frac{N(n, u)}{n}\right| < 2\epsilon,$$

and the existence of the limit (1.9) follows.

The next two steps of the proof are devoted to establishing (1.12). This asserts that the number of $m < n$ such that either

$$(\alpha) \quad H(m) < u \text{ and } H_k(m) > u$$

or

$$(\beta) \quad H(m) > u \text{ and } H_k(m) < u$$

is less than ϵn for each $k > k_0$, and sufficiently large n. It suffices (since the argument is the same for the other case) to consider only the case (α). At this point the *second* step of the proof comes in. It is proved that, given any $\delta > 0$, $\epsilon > 0$, for k fixed sufficiently large, and n sufficiently large,

$$|H(m) - H_k(m)| < \delta \tag{1.13}$$

except for at most $\frac{1}{2}\epsilon n$ integers $m < n$. Thus in case (α),

$$H(m) < u - \delta, \quad H_k(m) > u$$

can hold for at most $\frac{1}{2}\epsilon n$ integers $m < n$. Hence we need consider only those m for which

$$u - \delta < H(m) < u. \tag{1.14}$$

This then brings us to the *third* step of the proof. It is shown that given $\epsilon > 0$ there exists a $\delta > 0$ ($\delta = \delta(\epsilon)$, independent of u), such that for sufficiently large n, the number of $m < n$ such that (1.14) holds is less than $\frac{1}{2}\epsilon n$. This clearly completes the proof of the existence of $F(u)$. Furthermore, the result of this third step implies that for a fixed u, given any $\epsilon > 0$, there is a $\delta = \delta(\epsilon)$ such that $0 < F(u - \delta) - F(u) < \epsilon$, which yields the continuity of $F(u)$.

The main component of the argument used to carry out this last step is the result that, for any fixed integer l, the function

$$\Phi_l(x) = \frac{\phi(x)}{x} + \frac{\phi(x+1)}{x+1} + \dots + \frac{\phi(x+l)}{x+l}$$

has a continuous distribution function. Though we shall not bother to delineate the proof of this, it is contained in the arguments given. The idea in the proof of the result desired in the third step is that its negation would for some l imply the existence of a discontinuity in the distribution function of $\Phi_l(x)$.

2. First step: The existence of $F_k(u)$. We have

$$\begin{aligned}\sum_{n<x} \frac{\phi((n, A_k))}{(n, A_k)} &= \sum_{n<x} \sum_{d|(n,A_k)} \frac{\mu(d)}{d} \\ &= \sum_{d|A_k} \frac{\mu(d)}{d}\left[\frac{x}{d}\right] \\ &= x\sum_{d|A_k} \frac{\mu(d)}{d^2} - \sum_{d|A_k} \frac{\mu(d)}{d}\left\{\frac{x}{d}\right\},\end{aligned}$$

where $\{z\} = z - [z]$ denotes the fractional part of z. This in turn yields, from (1.10),

$$H_k(x) = -\sum_{d|A_k} \frac{\mu(d)}{d}\left\{\frac{x}{d}\right\}. \tag{2.1}$$

Since $\{x/d\}$ is, for fixed d, a periodic function of x with period d, we see from (2.1) that $H_k(x)$ is a periodic function of x with A_k as a period. Thus we have

$$N_k(n, u) = \sum_{\substack{m<n \\ H_k(m)>u}} 1 = \frac{n}{A_k} \sum_{\substack{m<A_k \\ H_k(m)>u}} 1 + O(1),$$

so that

$$\lim_{n\to\infty} \frac{N_k(n, u)}{n} = \frac{1}{A_k} N_k(A_k, u) = F_k(u)$$

exists.

3. Second step. We will prove in this section that, given any $\eta > 0$, for each $k > k_0 = k_0(\eta)$, and all $x > x_0 = x_0(k)$, we have

$$\sum_{n=1}^{x} (H(n) - H_k(n))^2 < \eta x. \tag{3.1}$$

From this it follows that if $M(x)$ is the number of $n < x$ such that

$$|H(n) - H_k(n)| > \delta,$$

$M(x) < \eta x/\delta^2$, which yields the statement concerning (1.13).

(3.1) is established in a rather straightforward fashion in the following sequence of lemmas.

LEMMA 3.1.

$$\sum_{n<x} H^2(n) \sim \left(\frac{1}{2\pi^2} + \frac{6}{\pi^4}\right) x. \tag{3.2}$$

Proof. This is essentially Lemma 12 of **(8)**, which asserts that

$$\int_1^x H^2(u)\, du \sim \frac{x}{2\pi^2}. \tag{3.3}$$

The passage from (3.3) to (3.2) is simple and we omit it. In passing it is perhaps of some interest to note that (3.3) is proved by means of a method of Walfisz **(2)**, and seems to be slightly "deeper" than the rest of our estimates which require only elementary methods together with a strong form of the prime number theorem.

LEMMA 3.2.

$$\sum_{n<x} H_k^2(n) \sim \alpha_k x, \tag{3.4}$$

where

$$\begin{aligned}\alpha_k = {} & \frac{1}{12} \sum_{\substack{d_1|A_k\\ d_2|A_k}} \frac{\mu(d_1)\,\mu(d_2)}{d_1^2 d_2^2} (d_1, d_2)^2 \\ & + \frac{1}{6} \prod_{p<p_k} \left(1 - \frac{1}{p^2}\right)^2 + \frac{1}{4} \prod_{p<p_k} \left(1 - \frac{1}{p}\right)^2 \\ & - \frac{1}{2} \prod_{p<p_k} \left(1 - \frac{1}{p}\right)\left(1 - \frac{1}{p^2}\right).\end{aligned} \tag{3.5}$$

Furthermore,

$$\lim_{k\to\infty} \alpha_k = \frac{1}{2\pi^2} + \frac{6}{\pi^4}. \tag{3.6}$$

Proof. From (2.1) we have

$$\begin{aligned}\sum_{n<x} H_k^2(n) &= \sum_{n<x} \sum_{\substack{d_1|A_k\\ d_2|A_k}} \frac{\mu(d_1)}{d_1} \frac{\mu(d_2)}{d_2} \left\{\frac{n}{d_1}\right\}\left\{\frac{n}{d_2}\right\} \\ &= \sum_{\substack{d_1|A_k\\ d_2|A_k}} \frac{\mu(d_1)\,\mu(d_2)}{d_1 d_2} \sum_{n<x} \left\{\frac{n}{d_1}\right\}\left\{\frac{n}{d_2}\right\}.\end{aligned} \tag{3.7}$$

Also,

$$\sum_{n<x}\left\{\frac{n}{d_1}\right\}\left\{\frac{n}{d_2}\right\} = \sum_{\substack{1\le i\le d_1-1\\ 1\le j\le d_2-1\\ (d_1,d_2)\mid(i-j)}} \frac{ij}{d_1d_2}\sum_{\substack{n<x\\ n\equiv i(d_1)\\ n\equiv j(d_2)}} 1 \tag{3.8}$$
$$= \frac{1}{d_1d_2}\sum_{\substack{1\le i\le d_1-1\\ 1\le j\le d_2-1\\ (d_1,d_2)\mid i-j}} ij\left(\frac{x}{\{d_1,d_2\}}+O(1)\right)$$
$$= \frac{\lambda}{d_1^2d_2^2}x\sum_{\substack{1\le i\le d_1-1\\ 1\le j\le d_2-1\\ \lambda\mid(i-j)}} ij + O(1),$$

where $\lambda = (d_1, d_2)$ is greatest common divisor of d_1, d_2, and $\{d_1, d_2\}$ is the least common multiple of d_1, d_2.

A simple calculation gives that

$$\sum_{\substack{1\le i\le d_1-1\\ 1\le j\le d_2-1\\ \lambda\mid i-j}} ij = \sum_{l=0}^{\lambda-1}\left(\sum_{\substack{i=1\\ i\equiv l(\lambda)}}^{d_1-1} i\right)\left(\sum_{\substack{j=1\\ j\equiv l(\lambda)}}^{d_2-1} j\right)$$
$$= \frac{d_1d_2}{\lambda}\left(\frac{\lambda^2}{12}+\frac{1}{6}+\frac{d_1d_2}{4}-\frac{(d_1+d_2)}{4}\right). \tag{3.9}$$

Combining (3.7), (3.8) and (3.9) yields

$$\sum_{n<x} H_k^2(n) = x\sum_{\substack{d_1\mid A_k\\ d_2\mid A_k}} \frac{\mu(d_1)}{d_1^2}\,\frac{\mu(d_2)}{d_2^2}\left\{\frac{\lambda^2}{12}+\frac{1}{6}+\frac{d_1d_2}{4}-\frac{d_1}{2}\right\} + O(1)$$
$$\sim x\left(\frac{1}{12}\sum_{\substack{d_1\mid A_k\\ d_2\mid A_k}} \frac{\mu(d_1)}{d_1^2}\,\frac{\mu(d_2)}{d_2^2}\,(d_1,d_2)^2\right.$$
$$+\frac{1}{6}\prod_{p<p_k}\left(1-\frac{1}{p^2}\right)^2+\frac{1}{4}\prod_{p<p_k}\left(1-\frac{1}{p}\right)^2$$
$$\left.-\frac{1}{2}\prod_{p<p_k}\left(1-\frac{1}{p}\right)\left(1-\frac{1}{p^2}\right)\right)$$

which is precisely (3.4) and (3.5). Since

$$\prod_{p<p_k}\left(1-\frac{1}{p}\right)\to 0 \quad\text{as } k\to\infty;\quad\text{and}\quad \prod_{p}\left(1-\frac{1}{p^2}\right)=\frac{6}{\pi^2},$$

it follows from (3.5) that

$$\lim_{k\to\infty}\alpha_k = \frac{1}{12}\sum_{d_1,d_2}\frac{\mu(d_1)\,\mu(d_2)\,(d_1,d_2)^2}{d_1^2d_2^2}+\frac{6}{\pi^2}.$$

(3.6) then follows from

$$\sum_{d_1,d_2}\frac{\mu(d_1)\,\mu(d_2)\,(d_1,d_2)^2}{d_1^2d_2^2} = \frac{6}{\pi^2}.$$

LEMMA 3.3.

$$\sum_{n<x} H(n)\,H_k(n) \sim \beta_k x \tag{3.10}$$

where

$$\beta_k = \frac{1}{12}\sum_{\substack{d_1|A_k\\ d_2|A_k}} \frac{\mu(d_1)}{d_1^2}\frac{\mu(d_2)}{d_2^2}(d_1, d_2)^2 - \frac{3}{2\pi^2}\prod_{p<p_k}\left(1-\frac{1}{p}\right) + \frac{1}{\pi^2}\prod_{p<p_k}\left(1-\frac{1}{p^2}\right). \tag{3.11}$$

Furthermore,

$$\lim_{k\to\infty}\beta_k = \frac{1}{2\pi^2} + \frac{6}{\pi^4}. \tag{3.12}$$

Proof. Setting

$$M(u) = \sum_{d<u}\frac{\mu(d)}{d},$$

since, by the prime number theorem, $M(u) = O(\log^{-c} u)$ for any fixed $c > 0$, we have for $uv = x$

$$\begin{aligned}\sum_{n<x}\frac{\phi(n)}{n} &= \sum_{dd'<x}\frac{\mu(d)}{d}\\ &= \sum_{d<u}\frac{\mu(d)}{d}\left[\frac{x}{d}\right] + \sum_{d'<v} M\left(\frac{x}{d'}\right) - M(u)[v]\\ &= \frac{6}{\pi^2}x - \sum_{d<u}\frac{\mu(d)}{d}\left\{\frac{x}{d}\right\} + O(v\log^{-2c} u).\end{aligned}$$

Taking $u = x\log^{-c} x$, we get

$$H(x) = -\sum_{d<x\log^{-c}x}\frac{\mu(d)}{d}\left\{\frac{x}{d}\right\} + O(\log^{-c} x). \tag{3.13}$$

(This is essentially Lemma 2 of (8).)

From (2.1) and (3.13) we obtain

$$\sum_{n<x} H_k(n)\,H(n) = \sum_{\substack{d_1|A_k\\ d_2<x\log^{-c}x}}\frac{\mu(d_1)}{d_1}\frac{\mu(d_2)}{d_2}\sum_{n<x}\left\{\frac{n}{d_1}\right\}\left\{\frac{n}{d_2}\right\} + O(\log^{-c} x).$$

Using a slight modification of (3.8) and (3.9) we get

$$\begin{aligned}\sum_{n<x} H_k(n)\,H(n) &= x\sum_{\substack{d_1|A_k\\ d_2<x\log^{-c}x}}\frac{\mu(d_1)}{d_1^2}\frac{\mu(d_2)}{d_2^2}\left\{\frac{\lambda^2}{12}+\frac{1}{6}+\frac{d_1d_2}{4}-\frac{(d_1+d_2)}{4}\right\}\\ &\qquad + O(x\log^{-c}x)\\ &\sim x\left(\frac{1}{12}\sum_{\substack{d_1|A_k\\ d_2|A_k}}\frac{\mu(d_1)}{d_1^2}\frac{\mu(d_2)}{d_2^2}(d_1,d_2)^2\right.\\ &\qquad\left. - \frac{3}{2\pi^2}\prod_{p<p_k}\left(1-\frac{1}{p_k}\right) + \frac{1}{\pi^2}\prod_{p<p_k}\left(1-\frac{1}{p^2}\right)\right)\end{aligned}$$

which gives (3.10) and (3.11). From this it follows easily that

$$\lim_{k\to\infty}\beta_k = \frac{1}{2\pi^2}+\frac{6}{\pi^4}.$$

Applying (3.2), (3.4), and (3.10) we have

$$\sum_{n<x} (H(n) - H_k(n))^2 = \sum_{n<x} H^2(n) - 2\sum_{n<x} H_k(n)\,H(n) + \sum_{n<x} H_k^2(n)$$
$$\sim x\left\{\frac{1}{2\pi^2} + \frac{6}{\pi^4} - 2\beta_k + \alpha_k\right\}.$$

From (3.6) and (3.12) we see that

$$\lim_{k\to\infty}\left(\frac{1}{2\pi^2} + \frac{6}{\pi^4} - 2\beta_k + \alpha_k\right) = 0,$$

and the assertion concerning (3.1) made at the beginning of this section follows immediately.

4. The third step. In this section we propose to prove that, given any $\epsilon > 0$, there exists a $\delta > 0$ such that the number of $m < x$ such that

(4.1) $$u < H(m) < u + \delta \qquad \text{for some } u,$$

is (for sufficiently large x) less than ϵx.

We shall suppose that the above statement is false and derive a contradiction. Negating the above assertion yields that for *some* constant $A > 0$ and *each* $\delta > 0$, there exist infinitely many positive integers x (depending possibly on δ) such that for some u (depending possibly on x as well as on u) the number of $m < x$ such that (4.1) is satisfied is at least Ax.

Since from (3.2) we have

(4.2) $$\sum_{m=1}^{x} H^2(m) < c_1 x,$$

it follows that for these $u = u(x, \delta)$ (we restrict ourselves to $0 < \delta < 1$), we have that either $-2 < u < 0$, or from (4.2)

$$\frac{u^2}{4} Ax < c_1 x,$$

so that in any event the possible values of $u = u(x, \delta)$ are bounded. Thus for each δ $(0 < \delta < 1)$ we can find an infinite sequence of positive integers $\{x_i(\delta)\}$ such that

(4.3) $$\lim_{i\to\infty} u(x_i(\delta), \delta) = u^*(\delta),$$

where the set of $u^*(\delta)$ is also clearly bounded. Thus again we can choose a sequence $\delta_j \to 0$ such that the limit

(4.4) $$\lim_{j\to\infty} u^*(\delta_j) = \bar{u}$$

exists.

Given any $\delta > 0$ we can find a $\delta_j < \frac{1}{3}\delta$ such that

$$|\bar{u} - u^*(\delta_j)| < \tfrac{1}{3}\delta.$$

Since from (4.3) we know that, for all sufficiently large i,

$$|u^*(\delta_j) - u(x_i(\delta_j), \delta_j)| < \tfrac{1}{10}\delta,$$

it follows that for this sequence $\{x_i(\delta_j)\}$ we have

$$|\bar{u} - u(x_i(\delta_j), \delta_j)| < \tfrac{2}{10}\delta. \tag{4.5}$$

For $m < x_i(\delta_j)$ there are more than Ax_i integers $m < x_i$ such that

$$u(x_i(\delta_j), \delta_j) < H(m) < u(x_i(\delta_j), \delta_j) + \delta_j. \tag{4.6}$$

But since (4.5) implies that

$$\bar{u} - \delta < u(x_i(\delta_j), \delta_j) < u(x_i(\delta_j), \delta_j) + \delta_j < \bar{u} + \delta,$$

it follows that for at least $\tfrac{1}{2}Ax_i$ of the $m < x_i$ we have one of the inequalities

$$(\alpha) \quad \bar{u} < H(m) < \bar{u} + \delta,$$
$$(\beta) \quad \bar{u} - \delta < H(m) < \bar{u}.$$

Since at least one of (α) or (β) must occur for a sequence of δ's approaching 0, at least one of these is the case for all $\delta > 0$. Since the treatment of the other case is exactly the same, we assume (α). Thus we have that, for any $\delta > 0$, there exist infinitely many positive integers n such that the number of integers $m < n$ for which

$$\bar{u} < \sum_{r=1}^{m} \frac{\phi(r)}{r} - \frac{6}{\pi^2} m < \bar{u} + \delta \tag{4.7}$$

is greater than $\tfrac{1}{2}An$.

Let $m_1 < m_2 < \ldots < m_t < n$ $(t > \tfrac{1}{2}An)$ be the integers $<n$ which satisfy (4.7). Clearly $m_{i+1} - m_i < 4/A$ has at least $\tfrac{1}{4}An$ solutions. Thus there exists an integer $l < 4/A$ such that $m_{i+1} - m_i = l$ has at least $A^2n/16$ solutions. Furthermore, by extracting a suitable subsequence from our infinite sequence of n, we may assume that l is independent of n.

The above in turn implies that for any $\delta > 0$ there exists an infinite sequence of n such that

$$\left| \sum_{r=m}^{m+l-1} \frac{\phi(r)}{r} - \frac{6}{\pi^2} l \right| < \delta \tag{4.8}$$

has at least $A^2n/16$ solutions $m < n$. In deriving a contradiction from this, the underlying idea is that this implies that the distribution function (it exists, though we forego a proof of this) of

$$\frac{\phi(x)}{x} + \ldots + \frac{\phi(x + l - 1)}{x + l - 1}$$

would have to have a discontinuity at $6l/\pi^2$, and this in turn would lead to the existence of a discontinuity in the distribution function of $\phi(x)/x$ (which is known to exist and be continuous) **(5)**.

We set

$$\frac{\phi_D(x)}{x} = \prod_{\substack{p|x \\ p<D}} \left(1 - \frac{1}{p}\right),$$

$$\mu_D(n) = \begin{cases} \mu(n) \text{ if } n \text{ is divisible only by primes } p < D, \\ 0 \quad \text{otherwise;} \end{cases}$$

so that

$$\frac{\phi_D(x)}{x} > \frac{\phi(x)}{x},$$

and

$$\begin{aligned} 0 < \sum_{n<x} \left\{\frac{\phi_D(n)}{n} - \frac{\phi(n)}{n}\right\} &= \sum_{d<x} \left\{\frac{\mu_D(d)}{d} - \frac{\mu(d)}{d}\right\}\left[\frac{x}{d}\right] \\ &= x \sum_{d<x} \frac{\mu_D(d) - \mu(d)}{d^2} + O(\log x) \\ &\sim x\left(\prod_{p<D} \left(1 - \frac{1}{p^2}\right) - \frac{6}{\pi^2}\right). \end{aligned}$$

From this it follows that, given $\eta_1, \eta_2 > 0$, we can choose $D > D(\eta_1, \eta_2)$ sufficiently large so that for all but $\eta_1 n$ integers $x < n$ we have

$$0 < \frac{\phi_D(x)}{x} - \frac{\phi(x)}{x} < \eta_2.$$

Thus, taking $\eta_2 = \delta/l$ and $\eta_1 = A^2/32$, we obtain from (4.8) that for each sufficiently large D, there exist infinitely many positive integers n such that the inequalities

$$\left|\sum_{r=m}^{m+l-1} \frac{\phi_D(r)}{r} - \frac{6}{\pi^2} l\right| < 2\delta \tag{4.9}$$

and

$$\left|\frac{\phi(m)}{m} + \sum_{r=m+1}^{m+l-1} \frac{\phi_D(r)}{r} - \frac{6}{\pi^2} l\right| < 2\delta \tag{4.10}$$

are satisfied simultaneously by at least $A^2n/32$ integers $m < n$.

LEMMA 4.1. *There exist absolute constants $\rho > 0$, and $\delta_0 > 0$ (independent of D) such that for at least $A^2n/64$ of the solutions $m < n$ of* (4.9) *and* (4.10) *we have for $\delta < \delta_0$*

$$\frac{6}{\pi^2} l - \sum_{r=m+1}^{m+l-1} \frac{\phi_D(r)}{r} > \rho. \tag{4.11}$$

Proof. For if (4.11) is false, (4.10) implies

$$0 < \frac{\phi(m)}{m} < \rho + 2\delta. \tag{4.12}$$

Since the distribution function of $\phi(m)/m$ exists and is continuous, for ρ and δ sufficiently small, (4.12) can have at most $A^2n/64$ solutions $m < n$.

Thus we may restrict ourselves to solutions m of (4.9) for which (4.11) holds. Also there is no loss of generality in assuming $\delta < \frac{1}{4}\rho$, as we shall do henceforth.

Next, we discard a certain "small" set of integers. Since

$$\sum_{m=1}^{n} \sum_{p \mid m} \frac{1}{p} = \sum_{p<n} \frac{1}{p}\left[\frac{n}{p}\right] < n \sum_{p} \frac{1}{p^2} = c_2 n \tag{4.13}$$

it follows that the number of $m < n$ such that

$$\sum_{p \mid m+i} \frac{1}{p} < E, \qquad 0 < i < l-1, \tag{4.14}$$

fails to hold is less than lc_2n/E, which for $E > 128lc_2/A^2$ is in turn less than $A^2n/128$. Thus for such an E we have an infinite sequence of n such that (4.9), (4.11) and (4.14) hold simultaneously for more than $A^2n/128$ integers $m < n$.

We now attempt to show that the set of integers m which satisfy (4.9), (4.11) and (4.14) has small density, thereby obtaining a contradiction. For a given integer m define

$$\lambda(m) = \prod_{\substack{p \mid m \\ p<D}} p.$$

We then associate with each integer m an $(l-1)$-dimensional vector $\vec{\lambda}(m)$ as follows:

$$\vec{\lambda}(m) = (\lambda(m+1), \lambda(m+2), \ldots, \lambda(m+l-1)).$$

Next, for a given vector $\vec{\lambda} = (\lambda_1, \ldots, \lambda_{l-1})$, wherein each λ_i is an integer which is a product of distinct primes $<D$, and

$$\sum_{p \mid \lambda_i} \frac{1}{p} < E, \qquad i = 1, \ldots, l-1, \tag{4.15}$$

and

$$\frac{6}{\pi^2} l - \sum_{i=1}^{l-1} \frac{\phi(\lambda_i)}{\lambda_i} > \rho, \tag{4.16}$$

we estimate the number of $m < n$ satisfying (4.9) such that $\vec{\lambda}(m) = \vec{\lambda}$ (possibly none). For such m we have

$$m + i \equiv 0 \pmod{\lambda_i}, \qquad i = 1, \ldots, l-1, \tag{4.17}$$

so that if there are any solutions they belong to a single arithmetic progression modulo $[\vec{\lambda}] = \{\lambda_1, \ldots, \lambda_{l-1}\}$, the least common multiple of the λ_i, $i = 1, \ldots, l-1$. Furthermore, in order that such solutions exist we must have

$$(\lambda_i, \lambda_j) \mid i - j \qquad i \neq j,\ 1 < i, j < l-1. \tag{4.18}$$

Suppose then that the aforementioned progression is $m \equiv \alpha \pmod{[\vec{\lambda}]}$. For those m such that $\vec{\lambda}(m) = \vec{\lambda}$ which satisfy (4.9) it follows that

$$\frac{6}{\pi^2} l - 2\delta - \sum_{i=1}^{l-1} \frac{\phi_D(\lambda_i)}{\lambda_i} < \frac{\phi_D(m)}{m} < \frac{6}{\pi^2} l + 2\delta - \sum_{i=1}^{l-1} \frac{\phi_D(\lambda_i)}{\lambda_i}, \tag{4.19}$$

so that for these m, $\phi_D(m)/m$ lies in a fixed interval of length 4δ which we shall denote by $I_\delta = I_\delta(\vec{\lambda})$. Thus the number of $m < n$ such that $\vec{\lambda}(m) = \vec{\lambda}$ and which satisfy (4.9) equals the number of $m < n$ which satisfy

(a) $m \equiv \alpha \pmod{[\vec{\lambda}]}$

(b) $\left(\dfrac{m+i}{\lambda_i}, \dfrac{\Delta}{\lambda_i}\right) = 1, \qquad i = 1, \ldots, l-1; \Delta = \prod_{p<D} p,$

(c) $\dfrac{\phi_D(m)}{m} \in I_\delta.$

LEMMA 4.2. *Given any $\eta > 0$, for D fixed sufficiently large, and δ sufficiently small (these requirements are however independent of $\vec{\lambda}$), the number of $m < n$ such that* (a), (b) *and* (c) *hold is less than*

$$(\eta n/[\vec{\lambda}]) \prod_{p<D} \left(1 - \frac{1}{p}\right)^{l-1}. \tag{4.20}$$

Proof. Suppose that the above statement concerning the estimate (4.20) is false, so that for infinitely many n, the number of $m < n$ satisfying (a), (b) and (c) is more than

$$(c_3 n/[\vec{\lambda}]) \prod_{p<D} \left(1 - \frac{1}{p}\right)^{l-1}. \tag{4.21}$$

Let $z_1, z_2, \ldots$ be those integers, composed of primes $<D$, which can occur as divisors of an integer $m \equiv \alpha \pmod{[\vec{\lambda}]}$ and such that (we denote the z_k generically by z)

$$\frac{\phi_D(z)}{z} = \frac{\phi(z)}{z} \in I_\delta. \tag{4.22}$$

From (4.16), (4.19) and our assumption $\delta < \rho/3$, (4.22) yields

$$\frac{\phi(z)}{z} > \frac{\rho}{3}. \tag{4.23}$$

Consider the number of $m < n$ such that (a), (b) above hold, and in addition for a fixed z,

$$m \equiv 0 \pmod{z}, \left(\frac{m}{z}, \Delta\right) = 1. \tag{d}$$

Clearly (d) implies (c).

Delete from $\Delta/[\vec{\lambda}]$ all prime factors $<l$ and any other prime factors of z; denoting the resulting integer by ψ. Then the number of $m < n$ which satisfy (a), (b), (d) is less than or equal to the number which satisfy (a) and

$$(m + i, \psi) = 1, \qquad i = 1, \ldots, l-1 \tag{b'}$$

and

$$m \equiv 0(z), \left(\frac{m}{z}, \psi\right) = 1. \tag{d'}$$

Setting $m = vz$ we have that the number of such m equals

$$(4.24)\qquad \sum_{\substack{v<n/z \\ (v,\psi)=1 \\ (vz+i)\equiv 0(\lambda_i) \\ (vz+i,\psi)=1 \\ 1\leq i\leq l-1}} 1 = \sum_{\substack{d_i \mid \psi \\ i=0,\dots,l-1}} \mu(d_0)\,\mu(d_1)\dots\mu(d_{l-1})\,. \sum_{\substack{v<n/z \\ v\equiv 0(d_0) \\ vz+i\equiv 0(d_i\lambda_i) \\ 1\leq i\leq l-1}} 1$$

Since $(d_i, z) = 1$, $(\lambda_i, \lambda_j) | i - j$, and the primes which divide ψ are $>l$, we see that the system of congruences

$$v \equiv 0(d_0), \quad vz + i \equiv 0(d_i\lambda_i), \qquad 1 \leq i \leq l - 1,$$

has solutions if and only if

$$(4.25)\qquad (d_i, d_j) = 1, \text{ and } (z, \lambda_i)|i; \qquad i \neq j, \quad 0 \leq i, \quad j \leq l - 1.$$

Furthermore, if (4.25) holds we have, since $(d_i, \lambda_j) = 1$,

$$\sum_{\substack{v<n/z \\ v\equiv 0(d_0) \\ vz+i\equiv 0(d_i\lambda_i) \\ 1\leq i\leq l-1}} 1 = \frac{n}{z} \frac{1}{\left\{d_0, \dfrac{d_1\lambda_1}{(z, \lambda_1)}, \dots, \dfrac{d_{l-1}, \lambda_{l-1}}{(z, \lambda_{l-1})}\right\}} + O(1)$$

$$= \frac{n}{zd_0d_1\dots d_{l-1}} \frac{1}{\left\{\dfrac{\lambda_1}{(z, \lambda_1)}, \dots, \dfrac{\lambda_{l-1}}{(z, \lambda_{l-1})}\right\}} + O(1).$$

Inserting this in (4.24) we get

$$(4.26)\qquad M = \frac{n}{z} \cdot \frac{1}{\left\{\dots, \dfrac{\lambda_i}{(z, \lambda_i)}, \dots\right\}} \sum_{\substack{d_i \mid \psi \\ (d_i, d_j)=1 \\ 0\leq i, j\leq l-1}} \frac{\mu(d_0)\dots\mu(d_{l-1})}{d_0\dots d_{l-1}} + O(1).$$

Since

$$\sum_{\substack{d_i \mid \psi \\ (d_i, d_j)=1 \\ 0\leq i, j\leq l-1}} \frac{\mu(d_0)\dots\mu(d_{l-1})}{d_0\dots d_{l-1}} = \sum_{c|\psi} \frac{\mu(c)}{c} l^{\nu(c)} = \prod_{p|\psi} \left(1 - \frac{l}{p}\right)$$

$$< c_4 \prod_{p|\psi} \left(1 - \frac{1}{p}\right)^l$$

$$< c_5 \left(\frac{z}{\phi(z)}\right)^l \prod_{p|[\vec{\lambda}]} \left(1 + \frac{1}{p}\right)^l \prod_{p<D} \left(1 - \frac{1}{p}\right)^l$$

and from (4.15), (4.23)

$$\prod_{p|[\vec{\lambda}]} \left(1 + \frac{1}{p}\right)^l < e^{El^{\nu}},$$

$$\left(\frac{z}{\phi(z)}\right)^l < \left(\frac{3}{\rho}\right)^l,$$

(4.26) yields (since $(z, \lambda_i)|i$)

$$M < c_6 \left(\frac{3}{\rho}\right)^l e^{El^{\nu}} l!\ (n/z[\vec{\lambda}]) \prod_{p<D} \left(1 - \frac{1}{p}\right)^l$$

or

$$(4.27)\qquad M < c_7(n/z[\vec{\lambda}]) \prod_{p<D} \left(1 - \frac{1}{p}\right)^l,$$

where $c_7 > 0$ is independent of D.

(4.27) together with (4.21) implies

$$\sum_k \frac{1}{z_k} \prod_{p<D}\left(1-\frac{1}{p}\right) > \frac{c_3}{c_7}. \tag{4.28}$$

On the other hand, the number of $m < n$ such that (d) holds for some z_k is, for large n, greater than or equal to

$$\frac{n}{2}\sum_k \frac{1}{z_k} \prod_{p<D}\left(1-\frac{1}{p}\right) > \frac{c_3}{2c_7}\, n,$$

using (4.28). Since for these m, $\phi_D(m)/m$ lies in a fixed interval I_δ of length 4δ, we see that for at least $c_3 n/4c_7$ of these m, $\phi(m)/m$ lies in a fixed interval of length 8δ (if D is large enough). Since $c_3/4c_7$ is independent of δ (and of D), this would contradict the continuity of the distribution function of $\phi(m)/m$. Thus the lemma is proved.

Finally then, letting T denote the number of $m < n$ which satisfy (4.9), (4.11) and (4.14), we have

$$T < \eta n \prod_{p<D}\left(1-\frac{1}{p}\right)^{l-1} \sum_{\vec{\lambda}} [\vec{\lambda}]^{-1}.$$

Since

$$\begin{aligned}\sum_{\vec{\lambda}} [\vec{\lambda}]^{-1} &< c_8\Big(\sum_{\lambda\mid\Delta} \frac{1}{\lambda}\Big)^{l-1}\\ &< c_8 \prod_{p<D}\left(1+\frac{1}{p}\right)^{l-1},\end{aligned}$$

we have

$$T < c_9 \eta n.$$

But for η sufficiently small, $c_9\eta < A^2/128$, so that we obtain a contradiction, and the proof is completed.

References

1. A. Walfisz, *On Euler's functions*, Akad. Nauk USSR, *90.4* (1953), 491–493.
2. ———, *Teilerprobleme* II, Math. Z., *34* (1931), 448–472.
3. S. S. Pillai and S. D. Chowla, *On the error term in some asymptotic formulae in the theory of numbers* I, J. London Math. Soc., *5* (1930), 95–101.
4. P. Erdös, H. N. Shapiro, *On the changes of sign of a certain error function*, Can. J. Math., *3* (1951), 375–383.
5. P. Erdös, *On the density of some sequences of numbers* III, J. London Math. Soc., *13* (1938), 119–127.
6. P. Erdös and A. Wintner, *Additive arithmetical functions and statistical independence*, Amer. J. Math., *61* (1939), 713–721.
7. S. Chowla and P. Erdös, *A theorem on the distribution of the values of L-functions*, J. Indian Math. Soc., *15* (1951), 11–18.
8. S. Chowla, *Contributions to the analytic theory of numbers*, Math. Z., *35* (1932), 280–299.

Notre Dame University *New York University*

ON EXPLICIT BOUNDS IN SCHOTTKY'S THEOREM

J. A. JENKINS

1. Introduction. To Schottky is due the theorem which states that a function $F(Z)$, regular and not taking the values 0 and 1 in $|Z| < 1$ and for which $F(0) = a_0$, is bounded in absolute value in $|Z| < r$, $0 < r < 1$, by a number depending only on a_0 and r. Let $K(a_0, r)$ denote the best possible bound in this result. Various authors have dealt with the problem of giving an explicit estimate for this bound. Qualitative estimates were given by Landau **(2)** and Valiron **(7)** and numerically evaluable estimates were given by Ostrowski **(4)**, Pfluger **(5)**, Ahlfors **(1)**, Robinson **(6)**, and Hayman **(3)**. In this paper we shall develop a simple method of obtaining such bounds. The bounds obtained are, in various situations, superior to any previously given. Our approach will be seen to be closest in spirit to that of Robinson.

2. Preliminary remarks. Our method depends initially on a simple application of the theory of subordination, a feature which it shares with most earlier methods (a notable exception being that of Ahlfors). Indeed it is clear that the maximum is attained for the function $F_0(Z)$ mapping $|Z| < 1$ onto the universal covering surface of the finite W-plane punctured at 0 and 1 and taking the value a_0 at $Z = 0$.

The function $F_0(Z)$ maps a domain D in $|Z| < 1$ bounded by three arcs of circles orthogonal to $|Z| = 1$ and with $Z = 0$ in its closure in a (1, 1) manner onto a half plane $\Im W > 0$ or $\Im W < 0$ with a_0 in its closure. Then a suitable branch of $w = \log F_0(Z) \mp \pi i$ maps $D(1, 1)$ onto the strip $-\pi < \Im w < 0$ or $0 < \Im w < \pi$ in these respective cases. Let $|Z| < 1$ be mapped onto $\Re z > 0$ in the following manner. In the first case let the sides of D corresponding to

$$\Im w = 0;\ \Im w = -\pi,\ \Re w > 0;\ \Im w = -\pi,\ \Re w < 0$$

go into the sets

$$\Im z = 0,\ \Re z > 0;\ \Im z = -\pi,\ \Re z > 0;\ |z - \tfrac{1}{2}\pi i| = \tfrac{1}{2}\pi,\ \Re z > 0.$$

In the second case let the sides of D corresponding to

$$\Im w = 0;\ \Im w = \pi,\ \Re w > 0;\ \Im w = \pi,\ \Re w < 0$$

go into the sets

$$\Im z = 0;\ \Re z > 0;\ \Im z = \pi,\ \Re z > 0;\ |z - \tfrac{1}{2}\pi i| = \tfrac{1}{2}\pi,\ \Re z > 0.$$

Denote the image of D by Δ. Let the mapping function be $Z = Z(z)$ and let us denote the function $F_0(Z(z))$ by $f(z)$. Evidently $\log f(z) \mp \pi i$ (with the choice of a suitable branch) maps the union $\mathfrak{D}$ of Δ, its reflection in the real

Received November 4, 1953.

axis and the positive real axis onto the strip $-\pi < \Im w < \pi$ and carries the semi-circular arc $|z| = \pi$, $\Re z > 0$ into the segment $\Re w = 0$, $-\pi < \Im w < \pi$.

Further let $\zeta = e^{-z}$, $\omega = e^{-w}$. It is verified at once that the composed mapping $\omega = \psi(\zeta)$ carries a subdomain E of $|\zeta| < 1$ having the origin as a centre of symmetry into the ω-plane slit along the negative real axis to the left of -1. We observe that $\psi(\zeta) = -(f(z))^{-1}$ where $\zeta = e^{-z}$. Let $\zeta = \phi(\omega)$ be the function inverse to ψ.

Let the point $z = b$ correspond to $Z = 0$. We will make use of the important remark due to Robinson that the maximum of $|F_0(Z)|$ on $|Z| = r$ is attained on the intersection of this circle with $\bar{D}$. The image of $|Z| \leqslant r$ lies in the half plane $\Re z \leqslant (1 + r)(1 - r)^{-1}\Re b$ so that $K(a_0, r)$ is not greater than the least upper bound of $|f(z)|$ when z lies in the intersection S of $\bar{\mathfrak{D}}$ and $\Re z \leqslant (1 + r)$ $(1 - r)^{-1}\Re b$.

With this formulation Hayman's Theorems I and III **(3)** reduce almost to trivialities. Indeed $|\psi(\zeta)/\zeta| \geqslant 1$ while the set $|\omega| < 1$ corresponds to a domain containing $|\zeta| < e^{-\pi}$, thus either $|\omega| \geqslant 1$ or $|\psi(\zeta)/\zeta| \leqslant e^{\pi}$. Thus for z in S

$$\log |f(z)| \leqslant \frac{1+r}{1-r}\Re b, \tag{1}$$

while if $|f(b)| \leqslant 1$, $\Re b \leqslant \pi$ and if $|f(b)| > 1$,

$$\Re b \leqslant \log |f(b)| + \pi.$$

Denoting $\mu = \max[1, |f(b)|]$ we have

$$\Re b \leqslant \log \mu + \pi; \tag{2}$$

combining this with (1) we get

$$\log |f(z)| \leqslant \frac{1+r}{1-r}(\log \mu + \pi),$$

that is,

$$K(a_0, r) \leqslant (\mu e^{\pi})^{(1+r)/(1-r)},$$

which is Hayman's Theorem I.

Let us impose the additional restriction $|f(b)| = e^{-\delta}$, $\delta > 0$. We verify at once that $(f(\pi^2/z))^{-1}$ has the same mapping properties as $f(z)$. Thus $\Re \pi^2 b^{-1} \geqslant \log |f(b)|^{-1} = \delta$. Evidently then $\Re b \leqslant \pi^2/\delta$ and by equation (1) for z in S

$$\log |f(z)| \leqslant \frac{\pi^2}{\delta}\frac{1+r}{1-r},$$

that is,

$$K(a_0, r) \leqslant \exp\left\{\frac{\pi^2}{\delta}\frac{1+r}{1-r}\right\},$$

which is Hayman's Theorem III.

The examples constituting Hayman's Theorems II and IV can be given just as simply in an explicit geometric manner. It should be remarked that at one point Hayman actually constructed mappings corresponding to ϕ and ψ above

but his proofs are greatly complicated by the unnecessary introduction of hyperbolic measure.

We can also obtain at once the asymptotic behaviour of $K(a_0, r)$ for fixed r as $|a_0|$ approaches ∞. Indeed, setting up the chain of mappings $U = V^2$, $\omega = 4V(1 - V)^{-2}$, $\zeta = \phi(\omega)$, $\xi = \zeta^2$, we see that the mapping from $|U| < 1$ to the ξ-plane is just that given by $\zeta = \phi(\omega)$ from $|\omega| < 1$ to the ζ-plane, i.e., $\xi = \phi(U)$ and

$$\phi(U) = \left\{\phi\left(\frac{4U^{\frac{1}{2}}}{(1-U^{\frac{1}{2}})^2}\right)\right\}^2.$$

Inserting the power series

$$\phi(U) = c_1 U + c_2 U^2 + \dots$$

in this functional equation, we can calculate all its coefficients. In particular $c_1 = 16c_1^2$, thus $c_1 = 1/16$. From this we derive readily the asymptotic behaviour due to Ostrowski **(4)**

$$K(a_0, r) \sim (16)^{2r/(1-r)}|a_0|^{(1+r)/(1-r)},$$

which, in fact, holds uniformly in r.

Now Robinson obtained not only the asymptotic behaviour but also universal bounds which imply the asymptotic result and which, indeed, for certain values of $|a_0|$ are better than Hayman's. In the next section we turn to the consideration of this question.

3. Bounds in Schottky's Theorem. We decompose the mapping $\omega = \psi(\zeta)$ into two stages, first mapping E by a function $\eta = \chi(\zeta)$ onto $|\eta| < 1$ so that $\chi(0) = 0$, $\chi'(0) > 0$ and $\chi(e^{-b}) = \eta_0$, say. Let the inverse function be $\zeta = \theta(\eta)$. Second we map $|\eta| < 1$ by the function $\omega = \lambda(\eta) = 4\eta(1 - \eta)^{-2}$ onto the ω-plane slit along the negative real axis to the left of -1. Clearly $\psi(\zeta) = \lambda(\chi(\zeta))$ and $\lambda(\eta_0) = -(f(b))^{-1}$. Note that $\lambda'(0) = 4$, $\chi'(0) = 4$.

Now, letting $|\eta| = r_1$,

$$|\lambda(\eta)| > \frac{4r_1}{(1 + r_1)^2},$$

thus

$$|\lambda(\eta)|^{-1} < \tfrac{1}{4}(r_1 + 2 + r_1^{-1}) \tag{3}$$
$$< \tfrac{1}{4}r_1^{-1} + \tfrac{3}{4}.$$

Combining this with the evident estimate $|\chi(\zeta)| > |\zeta|$ we have

$$|\psi(\zeta)|^{-1} < \tfrac{1}{4}|\zeta|^{-1} + \tfrac{3}{4},$$

i.e., for z in S

$$|f(z)| < \tfrac{1}{4}\,\{\exp(\Re b)\}^{(1+r)/(1-r)} + \tfrac{3}{4};$$

using the estimate (2) we obtain

$$K(a_0, r) < \tfrac{1}{4}\,(\mu e^{\pi})^{(1+r)/(1-r)} + \tfrac{3}{4},$$

a bound better than Hayman's in all cases.

This bound is not asymptotically best possible. To obtain such a bound we note that $\theta'(0) = \frac{1}{4}$ and that $\theta(\eta)$ is an odd function. Thus

$$|\theta(\eta)| < \frac{1}{4}\frac{r_1}{1 - r_1^2}$$

and

$$|\theta(\eta)|^{-1} > 4(r_1^{-1} - r_1).$$

Combining this with (3),

$$\begin{aligned} |\lambda(\eta)|^{-1} &< \tfrac{1}{4}(\tfrac{1}{4}|\theta(\eta)|^{-1} + 2r_1) + \tfrac{1}{2} \\ &< 16^{-1}|\theta(\eta)|^{-1} + 1; \end{aligned}$$

that is, for z in S,

$$|f(z)| < 16^{-1}\{\exp(\Re b)\}^{(1+r)/(1-r)} + 1. \tag{4}$$

We want now a corresponding bound for $\exp(\Re b)$. For this set $|\eta_0| = r_0$ and note

$$|\lambda(\eta_0)| < \frac{4r_0}{(1 - r_0)^2},$$

that is,

$$\frac{1 + r_0^2}{r_0} < 4|\lambda(\eta_0)|^{-1} + 2.$$

On the other hand

$$|\theta(\eta_0)| > \frac{1}{4}\frac{r_0}{1 + r_0^2},$$

so that

$$|\theta(\eta_0)|^{-1} < 4\frac{1 + r_0^2}{r_0}.$$

Hence

$$\exp(\Re b) < 16|f(b)| + 8.$$

Combining this with (4) we obtain

$$K(a_0, r) < 16^{-1}(16|a_0| + 8)^{(1+r)/(1-r)} + 1.$$

This bound is not only asymptotically best possible but in a universal bound of the form

$$K(a_0, r) < A(B|a_0| + C)^{(1+r)/(1-r)} + L,$$

the quantity B cannot be smaller than 16. If $B = 16$, A cannot be smaller than 1/16. If $B = 16$, $A = 1/16$, C cannot be smaller than $e^\pi - 16$ (between 7 and 8). In the estimate (4) the additive term 1 could not be replaced by anything smaller than 15/16.

We collect up these results and state:

THEOREM 1. *If $K(a_0, r)$ denotes the best possible bound in Schottky's Theorem for functions $F(Z)$ with $F(0) = a_0$ and if $\mu = \max\,[1, |a_0|]$ then*

$$K(a_0, r) < \tfrac{1}{4}\,(\mu e^\pi)^{(1+r)/(1-r)} + \tfrac{3}{4},$$

$$K(a_0, r) < \tfrac{1}{16}\,(16|a_0| + 8)^{(1+r)/(1-r)} + 1.$$

The second bound is much the better except for $|a_0|$ near 1 and r fairly large. It is clear that numerous other bounds could be obtained by combinations and modifications of the above estimates. Further, a small improvement of the constant 1 in (4) can be obtained by using the fact that $\theta(\eta)$ is a bounded function.

4. A bound in Landau's Theorem. Suppose that $F(Z)$ is regular for $|Z| < 1$ and does not take the values 0 and 1 and that

$$F(Z) = a_0 + a_1 Z + \ldots$$

is its Taylor expansion about $Z = 0$. Landau's Theorem states that $|a_1|$ has a bound depending only on a_0. Hayman **(3)** gave the explicit bound

$$|a_1| < 2|a_0|\{|\log|a_0|| + 5\pi\}.$$

We shall show that the constant 5π can be notably reduced using the approach presented above.

We observe first that obtaining a bound of the form

$$|a_1| < 2|a_0|\{|\log|a_0|| + K\}$$

it is enough to confine ourselves to the case $|a_0| \geqslant 1$, $|a_0 - 1| \geqslant 1$. Indeed, if it is proved in this case and we have $|F(0)| \geqslant 1$, $|F(0) - 1| < 1$, we consider the function $\Phi(Z) = F(Z)/(F(Z) - 1)$. For it

$$|\Phi(0)| = |F(0)/(F(0) - 1)| \geqslant 1,$$
$$|\Phi(0) - 1| = |F(0) - 1|^{-1} \geqslant 1;$$

thus

$$|\Phi'(0)| = |a_1/(a_0 - 1)^2| < 2|a_0/(a_0 - 1)|\{|\log|a_0/(a_0 - 1)|| + K\},$$

that is,

$$\begin{aligned} |a_1| &< 2|a_0||a_0 - 1|\{|\log|a_0/(a_0 - 1)|| + K\} \\ &< 2|a_0|\{|\log|a_0|| + K - |a_0 - 1|\log|a_0 - 1| - K(1 - |a_0 - 1|)\}. \end{aligned}$$

Since, for $|a_0 - 1| < 1$,

$$|a_0 - 1|\log|a_0 - 1| + K(1 - |a_0 - 1|) \geqslant 0$$

provided $K \geqslant 1$, which is necessarily so in the present situation, the result is true for all $|a_0| \geqslant 1$. Applying a similar argument to the function $\Psi(Z) = (F(Z))^{-1}$ in the case $|a_0| < 1$ we obtain the same bound for all $|a_0|$.

Now for the function $\phi(\omega)$ we verify by an elementary calculation

$$\phi'(-a_0^{-1}) = -\frac{2a_0^2\Re b}{a_1 e^b}.$$

Further we have for $|\omega| < 1$ the well-known estimate

$$\left|\frac{1}{\zeta}\frac{d\zeta}{d\omega}\right| \geqslant \frac{1}{|\omega|}\frac{1 - |\omega|}{1 + |\omega|}.$$

Applying this at the value $\omega = -a_0^{-1}$ we derive

$$|a_1| < \frac{1 + |a_0|^{-1}}{1 - |a_0|^{-1}} 2|a_0|\Re b.$$

It is clear that this will lead to an advantageous bound for $|a_0|$ large. In particular, if $|a_0|^{-1} < \epsilon(1+\epsilon)^{-1}$,

$$\begin{aligned} |a_1| &< (1 + 2(1+\epsilon)|a_0|^{-1})2|a_0|\Re b \\ &< 2|a_0|\Re b + 4(1+\epsilon)\Re b. \end{aligned}$$

We recall

$$\Re b < \log(16\,|a_0| + 8).$$

Thus for $|a_0| > t$ we have

$$\begin{aligned} |a_1| < 2|a_0|\Big\{&\log|a_0| + \log 16 + \log\left(1 + \frac{1}{2t}\right) \\ &+ \frac{2t}{t-1}\left(e^{-1} + \frac{1}{t}\left[\log 16 + \log\left(1 + \frac{1}{2t}\right)\right]\right)\Big\}, \end{aligned}$$

using the facts that $x^{-1}\log x < e^{-1}$ and $1 < |a_0|/t$ for $|a_0| > t$. In particular for $|a_0| > 2.63$

$$|a_1| < 2|a_0|\{\log|a_0| + 7.77\}. \tag{5}$$

It remains to find a bound advantageous when $|a_0|$ is near 1. Originally we used for this an improved version of Hayman's technique. However we now use an improved version of a suggestion by the referee which gives a somewhat simpler argument and also a slightly better value of the final constant.

We first make the trivial observation that if a domain D_1 in the z-plane is mapped conformally into a domain D_2 in the w-plane by a function $w = w(z)$ with the point z_0 going into the point w_0 and if the inner radius of D_1, D_2 with respect to z_0, w_0 is r_1, r_2 then $|w'(z_0)| < r_2/r_1$. Now the mapping constructed above from the z-plane to the w-plane corresponds to the maximal value of $|a_1|$ for given a_0, by the majoration principle. Under it the half-plane $\Re z > \frac{1}{2}\pi$ is mapped into the w-plane slit along the half-infinite segments $\Im w = (2n+1)\pi$, $\Re w < 0$, n running through all integers. Under the assumptions $|a_0| > 1$, $|a_0 - 1| > 1$ we have $\Re b > 3^{\frac{1}{2}}\pi/2$. The inner radius of $\Re z > \pi/2$ with respect to the point b is $2(\Re b - \pi/2)$. The inner radius of the w-plane slit as above with respect to the image of b (namely $\log a_0 \mp \pi i$) is $2|1 - a_0^{-1}|^{\frac{1}{2}} \log |2a_0 - 1 + 2(a_0(a_0-1))^{\frac{1}{2}}|$, the radicals being properly determined. The derivative of the mapping function at the point $z = b$ is $a_1/2a_0\Re b$. Thus

$$|a_1| < 2(|a_0||a_0 - 1|)^{\frac{1}{2}} \log|2a_0 - 1 + 2\{a_0(a_0-1)\}^{\frac{1}{2}}|\Re b/(\Re b - \pi/2).$$

Since $\Re b > 3^{\frac{1}{2}}\pi/2$ we have for $|a_0| = t$, $1 < t$,

$$\begin{aligned} |a_1| &< (3 + 3^{\frac{1}{2}})|a_0|(1 + t^{-1})^{\frac{1}{2}}\{\log|a_0| + \log[2 + t^{-1} + 2(1+t^{-1})^{\frac{1}{2}}]\} \\ &< 2|a_0|\{\log|a_0| + L(t)\}, \end{aligned}$$

where

$$L(t) = \tfrac{1}{2}(3 + 3^{\frac{1}{2}})(1 + t^{-1})^{\frac{1}{2}} \log[2t^2 + t + 2t(t^2 + t)^{\frac{1}{2}}] - 2\log t.$$

A direct calculation verifies that $L(t)$ is an increasing function of t for $t > 1$. Thus for $1 < |a_0| < t$ we have

$$|a_1| < 2|a_0|\{\log|a_0| + L(t)\}.$$

In particular, for $1 < |a_0| < 2.63$,

$$|a_1| < 2|a_0|\{\log|a_0| + 7.77\}.$$

Combining this with (5) and our earlier remarks we have

THEOREM 2. *If $F(Z)$ is regular for $|Z| < 1$, does not take the values 0 and 1 and has Taylor expansion about $Z = 0$,*

$$F(Z) = a_0 + a_1 Z + \ldots,$$

then

$$|a_1| < 2|a_0|\{|\log|a_0|| + 7.77\}.$$

It should be observed that in an expression of this form the number 2 could not be replaced by any smaller number. We have replaced Hayman's constant 5π by 7.77 which is less than half as big. Our earlier method gave the constant 8.58.

REFERENCES

1. L. V. Ahlfors, *An extension of Schwarz's Lemma*, Trans. Amer. Math. Soc., *43* (1938), 359–364.
2. H. Bohr and E. Landau, *Über das Verhalten von $\zeta(s)$ und $\zeta_\kappa(s)$ in der Nähe der Geraden $\sigma = 1$*, Nach. Akad. Wiss. Göttingen, Math.-Phys. Kl., 1910, 303–330.
3. W. K. Hayman, *Some remarks on Schottky's Theorem*, Proc. Cambridge Phil. Soc., *43* (1947), 442–454.
4. A. Ostrowski, *Asymptotische Abschätzung des absoluten Betrages einer Funktion, die die Werte 0 und 1 nicht annimmt*, Comm. Math. Helv., *5* (1933), 55–87.
5. A. Pfluger, *Über numerische Schranken im Schottky'schen Satz*, Comm. Math. Helv., *7* (1934–35), 159–170.
6. R. M. Robinson, *On numerical bounds in Schottky's Theorem*, Bull. Amer. Math. Soc., *45* (1939), 907–910.
7. G. Valiron, *Compléments au théorème de Picard-Julia*, Bull. Sc. Math., *51* (1927), 167–183.

University of Notre Dame

ON REGIONS OMITTED BY UNIVALENT FUNCTIONS II

A. W. GOODMAN AND E. REICH

1. Introduction. Let S denote the family of functions $f(z)$ regular and univalent in $|z| < 1$, with the expansion $f(z) = z + a_2 z^2 + \ldots$ about $z = 0$, and let A_f denote the area of the intersection of the open circle $|w| < 1$ with D_f, the image of $|z| < 1$ under $f(z)$. A few years ago one of the authors **(1)** proved that if

$$A = \inf_{f \in S} \{A_f\} \tag{1}$$

then

$$.5000\pi < A < .7728\pi. \tag{2}$$

Recently Jenkins **(2)**, by a rather ingenious application of circular symmetrization, improved the lower bound for A to

$$.5387\pi < A. \tag{3}$$

By introducing only a slight change in Jenkins' work we obtain a still better lower bound, namely

$$.62\pi < A, \tag{4}$$

and at the same time we generalize Jenkins' result on the maximal length of arc on a circle $|w| = r$ left uncovered by D_f.

2. Values omitted by functions in $S(s)$. This section is devoted to proving the generalization just mentioned. Let s be the distance from the origin to the boundary of D_f. It is well known that $\frac{1}{4} \leqslant s \leqslant 1$ and that the extreme values correspond to the unique functions $f(z) = z/(1 + e^{i\theta}z)^2$, and $f(z) = z$, respectively. If we form the set $S(s)$ of all functions of S corresponding to a given fixed s, then as s runs from $\frac{1}{4}$ to 1 the sets $S(s)$ will exhaust S.

THEOREM 1. *Let $f(z) \in S(s)$, $\frac{1}{4} \leqslant s \leqslant 1$ and let $L(r, s)$ be the length of arc of the circle $|w| = r$ not covered by D_f. If*

$$s \leqslant r \leqslant s/(2s^{\frac{1}{2}} - 1) \tag{6}$$

then

$$L(r, s) \leqslant 2r \cos^{-1}(4s^{\frac{1}{2}} - 1 - r(8 - 4s^{-\frac{1}{2}})) \equiv \phi(r, s) \tag{7}$$

and this inequality is sharp in both variables.

Received February 23, 1954. The results presented in this paper were obtained independently and almost simultaneously by the two authors. Their results were identical up to and including equation (25), where it becomes necessary to estimate the maximum of the transcendental function defined by equations (24) and (25). Goodman obtained $A > .6028\pi$, while Reich by a suitable transformation, equation (27), and considerably more computation achieved $A > .62\pi$. In the case of Reich the work forms part of a Ph.D. thesis supervised by Professor E. F. Beckenbach at the University of California, Los Angeles.

Remarks. It is clear that if $r < s$, then $L(r, s) = 0$, and it will develop in the proof that if $r > s/(2s^{\frac{1}{2}} - 1)$ the entire circle $|w| = r$ may be omitted. In the range (6) the function

$$a(r, s) \equiv 4s^{\frac{1}{2}} - 1 - r(8 - 4s^{-\frac{1}{2}}) \tag{8}$$

satisfies $|a(r, s)| < 1$, and the principal branch of $\cos^{-1}$ is to be used in (7). In the special case $r = s$, (7) becomes

$$L(r, r) < 2r \cos^{-1}(8r^{\frac{1}{2}} - 8r - 1), \tag{9}$$

a result obtained by Jenkins. The bound (9) is valid for all functions of S, but equality can occur only if $r = s$, since $\phi(r, s)$ is an increasing function of s for each r in $\frac{1}{4} < s < r < 1$.

Proof. We begin by constructing an explicit conformal mapping. Consider, in the ζ-plane, the domain D bounded by $|\zeta| = 1$ together with the two portions of the real axis $\zeta_1 = -\rho_1 < \zeta < -1$ and $1 < \zeta < \infty$, and distinguish the point $\zeta_0 = -\rho_0$ $(\rho_0 > \rho_1 > 1)$ within this domain. The function $\eta = \zeta + \zeta^{-1} + 2$ maps this domain on the η-plane slit along the real axis from $\eta_1 = -(\rho_1 + \rho_1^{-1} - 2) < 0$ to $+\infty$. The point ζ_0 goes into $\eta_0 = -(\rho_0 + \rho_0^{-1} - 2) < \eta_1$. The function $W = (\eta - \eta_1)^{\frac{1}{2}}$ (taking the positive determination on the upper side of the positive real axis) maps the preceding domain on the upper half W-plane, η_0 going into $W_0 = i(\eta_1 - \eta_0)^{\frac{1}{2}}$. Finally $z = -(W - W_0)/(W - \bar{W}_0)$ maps the latter domain on $|z| < 1$, W_0 going into $z = 0$. An elementary calculation shows that

$$\left.\frac{d\zeta}{dz}\right|_{z=0} = \frac{4\rho_0(\rho_0 - 1)}{\rho_0 + 1} - \frac{4\rho_0^2(\rho_1 - 1)^2}{\rho_1(\rho_0^2 - 1)} > 0. \tag{10}$$

On the other hand the function $Z = \rho_0(\zeta^2 + \rho_0\zeta)/(\rho_0\zeta + 1)$ maps D on a domain D^* bounded by an arc on $|Z| = \rho_0$ placed symmetrically with respect to the positive real axis, together with the portions

$$Z_1 = \frac{\rho_0\rho_1(\rho_0 - \rho_1)}{\rho_0\rho_1 - 1} < Z < \infty$$

of the latter. The point ζ_0 goes into $Z = 0$. By an elementary calculation

$$\left.\frac{dZ}{d\zeta}\right|_{\zeta=\zeta_0} = \rho_0^2(\rho_0^2 - 1).$$

As a function of z, Z maps $|z| < 1$ on the domain D^* and

$$\left.\frac{dZ}{dz}\right|_{z=0} = \frac{4\rho_0^3}{(\rho_0 + 1)^2} - \frac{4\rho_0^4(\rho_1 - 1)^2}{\rho_1(\rho_0^2 - 1)^2} \equiv K. \tag{11}$$

Thus the function $w = Z/K$ belongs to $S(s)$, where

$$s = \frac{Z_1}{K} = \frac{\rho_1^2(\rho_0^2 - 1)^2}{4\rho_0^2(\rho_0\rho_1 - 1)^2}. \tag{12}$$

It maps $|z| < 1$ onto a domain $\bar{D}(r, s)$ bounded by the portion $s < z < \infty$ of the real axis and an arc of the circle $|w| = r$, placed symmetrically with respect to the real axis. An easy computation yields

$$r = \frac{\rho_0}{K} = \frac{\rho_1(\rho_0^2 - 1)^2}{4\rho_0^2(\rho_0\rho_1 - 1)(\rho_0 - \rho_1)}. \tag{13}$$

To determine the length of arc on $|w| = r$ on the boundary of $\bar{D}(r, s)$ we note that the end points of the arc in D^* are the images of the points on $|\zeta| = 1$ where $dZ/d\zeta = 0$, i.e., the solutions of $\rho_0\zeta^2 + 2\zeta + \rho_0 = 0$. These are the points $\zeta = (-1 \pm i(\rho_0^2 - 1)^{\frac{1}{2}})/\rho_0$ and their images are

$$Z = (\rho_0^2 - 2 \pm 2i(\rho_0^2 - 1)^{\frac{1}{2}})/\rho_0.$$

The angle subtended by this arc is $\theta = 2\cos^{-1}(1 - 2/\rho_0^2)$ (the principal branch of $\cos^{-1}$ being used). It remains to determine ρ_0 as a function of s. Solving equations (12) and (13) simultaneously, yields

$$\rho_0 = \frac{s^{1/4}}{((2r - s^{\frac{1}{2}})(2s^{\frac{1}{2}} - 1))^{\frac{1}{2}}}, \tag{14}$$

and

$$\rho_1 = \frac{s^{3/4}}{r}\left(\frac{2r - s^{\frac{1}{2}}}{2s^{\frac{1}{2}} - 1}\right)^{\frac{1}{2}}. \tag{15}$$

The conditions $\frac{1}{4} < s < 1$, $s < r$ required by the domain $\bar{D}(r, s)$, imply that $2r - s^{\frac{1}{2}} > 0$ and $2s^{\frac{1}{2}} - 1 > 0$, so that (14) and (15) determine positive values of ρ_0 and ρ_1. The construction of a function mapping $|z| < 1$ onto $\bar{D}(r, s)$ can be effected in the manner described if ρ_0 and ρ_1 determined by (14) and (15) satisfy

$$\rho_0 > \rho_1 > 1. \tag{16}$$

The condition $\rho_1 > 1$ leads from (15) to the equivalent condition

$$(r - s)(r + s - 2rs^{\frac{1}{2}}) > 0, \tag{17}$$

and the requirement that $\rho_0 > \rho_1$ yields

$$\rho_0 - \rho_1 = \frac{s^{1/4}(r + s - 2rs^{\frac{1}{2}})}{r((2r - s^{\frac{1}{2}})(2s^{\frac{1}{2}} - 1))^{\frac{1}{2}}} > 0, \tag{18}$$

and since the other factors are positive, both (17) and (18) are satisfied if and only if

$$r + s - 2rs^{\frac{1}{2}} > 0. \tag{19}$$

This yields the right side of (6) and also gives

$$s > 2r^2 - r - 2r(r^2 - r)^{\frac{1}{2}}. \tag{20}$$

This last expression coincides with the lower bound given in a theorem due to Pick and Nevanlinna (proved again in **(1)**) whereby a bounded function

$|g(z)| < r$, $g(z) \in S$, cannot omit a point with modulus less than the left side of (20). This shows that if $r > s/(2s^{\frac{1}{2}} - 1)$ then the entire circle $|w| = r$ may be omitted by $f(z)$.

We have thus shown that if r and s are given, subject to the restriction (6), and $\frac{1}{4} < s < 1$, then (14) and (15) determine ρ_0 and ρ_1 satisfying (16), and the function $f(z)$ mapping $|z| < 1$ onto $\tilde{D}(r, s)$ can be constructed as described. For this function the arc omitted on $|w| = r$ has length

$$2r \cos^{-1}(1 - 2/\rho_0^2)$$

and using (14) this yields $\phi(r, s)$, the left side of (7).

The proof that $\phi(r, s)$ is a maximum for $L(r, s)$ in the class of functions $S(s)$ is almost identical with that given by Jenkins **(2)**, in the special case $r = s$, and hence will not be reproduced here.

3. A lower bound for A. We now apply Theorem 1 to obtain a new lower bound for A. We consider the various subsets $S(s)$ of S and (with the notation of §1) set

$$A(s) = \inf_{f \in S(s)} \{A_f\}, \tag{21}$$

and hence

$$A = \inf_{\frac{1}{4} < s < 1} \{A(s)\}. \tag{22}$$

Let $B(s) = \pi - A(s)$, then by Theorem 1

$$B(s) < B^*(s) \equiv \int_s^1 \phi(r, s)\, dr. \tag{23}$$

This integral can be evaluated in terms of the elementary functions:

$$B^*(s) = F(\alpha, \beta, 1) - F(\alpha, \beta, s) \tag{24}$$

where

$$F(\alpha, \beta, r) = \alpha^{-2}\{(\alpha^2 r^2 - \beta^2 - \tfrac{1}{2}) \cos^{-1}(\alpha r + \beta) + \tfrac{1}{2}(3\beta - \alpha r)(1 - (\alpha r + \beta)^2)^{\frac{1}{2}}\}, \tag{25}$$

where $\alpha = 4s^{-\frac{1}{2}} - 8$ and $\beta = 4s^{\frac{1}{2}} - 1$. All that remains is to determine the maximum value of $B^*(s)$ for s in the interval $(\frac{1}{4}, 1)$. This, unfortunately, leads to a transcendental equation and we are forced to make a dull and detailed set of computations to secure a valid numerical bound for $B^*(s)$.

A table of $B^*(s)$ partially reproduced at the end of the paper, seems to indicate that $B^*(s) < 1.18 < .376\pi$, which would imply (4). To prove (4) it is however sufficient to show that

$$B^*(s) < 1.19 < .38\pi. \tag{26}$$

The proof of (26) involved a great deal of tedious computation, but, briefly, was accomplished as follows.

In (23) change the variable of integration from r to t where

$$t^2 = 1 - 2s^{\frac{1}{2}} + r(4 - 2s^{-\frac{1}{2}}) = \tfrac{1}{2}\{1 - a(r, s)\} \geq 0.$$

Also put $\xi = 2s^{\frac{1}{2}} - 1$. As a result of these transformations we can write

$$B^*(s) = G(\xi) = \frac{(1+\xi)^2}{2\xi^2} \int_\xi^{\gamma(\xi)} (\xi + t^2)\, t \sin^{-1} t \, dt \tag{27}$$

where $0 < \xi \leq 1$ for $\frac{1}{4} < s \leq 1$, and

$$\gamma(\xi) = \left(\frac{3\xi - \xi^2}{1+\xi}\right)^{\frac{1}{2}}.$$

The formula (27) was found convenient for estimation because of the convexity properties of $\gamma(\xi)$ and the inverse sine. The next step was to find bounds for $\Delta G(\xi) = G(\xi + \Delta\xi) - G(\xi)$ as a function of ξ and $\Delta\xi$, useful for small $|\Delta\xi|$. This is a straightforward task, starting from (27), but we do not reproduce the result here as it is quite complicated and uninteresting. The last step of the procedure was to evaluate $G(\xi)$ at a series of about forty unequally spaced mesh points $\{\xi_i\}$ in $[0, 1]$, and to apply the upper bound that had been obtained for $|\Delta G(\xi)|$ to show that $G(\xi)$ did not exceed 1.19 in any interval (ξ_i, ξ_{i+1}).

4. A supplementary remark. The upper bound on A stated in (2) shows that there exist $f \in S$ which omit an area

$$\pi - A_f > .7137 \tag{28}$$

in $|w| < 1$. Let $\bar{S}$ be the subset of S for which (28) holds. We now prove

THEOREM 2. *$f \in \bar{S}$ implies that D_f contains the circle $|w| \leq .295$.*

Proof. We show that $G(\xi) < 0.7135$ for $0 < \xi < 0.0868$.

Since

$$\sin^{-1} t \leq \frac{\sin^{-1}\gamma(\xi)}{\gamma(\xi)}\, t, \qquad 0 \leq t \leq \gamma(\xi),$$

we have, using (27),

$$G(\xi) \leq \frac{(1+\xi)^2 \sin^{-1}\gamma(\xi)}{2\xi^2\gamma(\xi)} \int_0^{\gamma(\xi)} (\xi t^2 + t^4)\, dt = \frac{(7+\xi)(3-\xi)\sin^{-1}\gamma(\xi)}{15}$$

$$\leq \frac{7+\xi}{5} \sin^{-1}\gamma(\xi) < 0.7135,$$

if $\xi \leq 0.0868$.

COROLLARY. *If there exists a function $g \in S$ for which $A_g = A$, then D_g contains the circle $|w| \leq 0.295$.*

s	$B^*(s)/\pi$	$A^*(s)/\pi \equiv 1-B^*(s)/\pi$
.25	.00000	1.00000
.30	.22639	.77361
.35	.29716	.70284
.40	.33799	.66201
.45	.36164	.63863
.50	.37305	.62695
.55	.37457	.62543
.60	.36741	.63259
.65	.35218	.64782
.70	.32908	.67092
.75	.29806	.70194
.80	.25881	.74119
.85	.21081	.78919
.90	.15315	.84685
.95	.08425	.91574
1.00	.00000	1.00000

References

1. A. W. Goodman, *Note on regions omitted by univalent functions*, Bull. Amer. Math. Soc., *55* (1949), 363–369.
2. J. A. Jenkins, *On values omitted by univalent functions*, Amer. J. Math., *75* (1953), 406–408.

University of Kentucky *Institute for Advanced Study*

CONVOLUTION TRANSFORMS RELATED TO NON-HARMONIC FOURIER SERIES

D. B. SUMNER

1. Introduction. Widder has pointed out (**2**, p. 219) in connection with Wiener's fundamental work on the operational calculus (**1**, pp. 557–584), that the convolution transform

$$f(x) = \int_{-\infty}^{\infty} G(x-t)\,\phi(t)\,dt \tag{1.1}$$

will be inverted by the operator $DE(D)$, where $D = d/dx$, and

$$1/wE(w) = \int_{-\infty}^{\infty} \exp(-xw)\,G(x)\,dx,$$

where a suitable interpretation must be found for $E(D)$. Cases where $E(w)$ is entire have been considered by Widder (**2**, pp. 217–249; **3**, pp. 7–60), Hirschman and Widder (**4**, pp. 659–696; **8**, pp. 135–201), and the author (**5**).

The most general method of interpreting $E(D)$ is as $\lim_{n\to\infty} P_n(D)$, where $P_n(w)$ is a polynomial of degree n, the method requiring a knowledge of $f(x)$ only for real values of its argument. However in cases where more is known about $E(w)$ (**4**, p. 692; **5**, pp. 174–183; **6**, p. 219), it is possible to represent $E(w)$ as an integral, when the computations are simpler, but it is necessary to have $f(x)$ defined for complex arguments.

The purpose of this article is to consider convolution transforms for which the invertor function $E(w)$ is entire, is not necessarily even, and can be represented by a Fourier-Lebesgue integral. The real numbers which are taken to be the zeros of $E(w)$ are a generalization of the non-harmonic Fourier exponents discussed by Levinson (**7**, pp. 47–57). The classical Stieltjes transform (**3**), and the generalized form of it (**5**), are particular cases. The assumptions made about the zeros of $E(w)$ are sufficient to establish all properties needed, and no integrability condition is postulated for $E(u)$.

2. Definitions. We suppose throughout that

$$\lambda_n = \rho + n - \delta + 2\delta\alpha_n,\ \mu_n = n - \delta + 2\delta\beta_n \qquad (n = 1, 2, \ldots), \tag{2.1}$$
$$0 \leqslant \alpha_n, \beta_n \leqslant 1,\quad 0 < \delta < \tfrac{1}{4},\quad 0 < \rho < 1 - 2\delta;$$

$$E(w) = \prod_{1}^{\infty} (1 - w/\lambda_n)(1 + w/\mu_n); \tag{2.2}$$

$$G(z) = \lim_{R\to\infty} (2\pi i)^{-1} \int_{c-iR}^{c+iR} \exp(zw)\,dw/wE(w), \qquad 0 < c < \lambda_1. \tag{2.3}$$

The symbols A, A_k denote absolute constants throughout.

Received September 12, 1952; in revised form July 15, 1954.

3. Some properties of $E(w)$. The numbers λ_n, μ_n are those used by Levinson (7, pp. 47–57) in his work on non-harmonic Fourier series. With the notation $w = u + iv = r \exp(i\phi)$, Levinson's methods may be used to establish the following inequalities:

(3.1) $$|E(w)| < A \exp(\pi|v|)/r^{\rho+1-4\delta}, \qquad (r > 1);$$

(3.2) $$|E(w)| > A \exp(\pi|v|)/r^{\rho+1-4\delta}, \text{ provided that}$$

(3.21) $r > 1$, $|w - r_n| > \Delta > 0$, $r_n = \lambda_n$ *or* $-\mu_n$;

(3.3) $$|E'(r_n)|^{-1} < Ar^{\rho+1+4\delta};$$

(3.4) *there exists a constant* q, $1 < q < 2$, *such that* $E(u) \in L^q(-\infty, \infty)$.

For the behaviour of $E(w)$ along the imaginary axis, we establish the more precise inequalities

(3.5) $$|E(iv)| < A \exp(\pi|v|)/|v|^{\rho+1-2\delta}, \; |E(iv)| > A \exp(\pi|v|)/|v|^{\rho+1+2\delta};$$

(3.6) $$|\text{amp } E(iv\theta)/E(iv)| < A(1 - \theta),$$

where $0 < \theta < 1$, *and the constant is independent of* v;

(3.7) $|E(iv\theta)/E(iv)|$ *is a decreasing function of* $|v|$, $0 < \theta < 1$.

Proof of (3.1). Let $\Re(w) > 0$, and N be the integer defined by

(3.8) $$(\rho + N - \tfrac{1}{2}) \cos\phi < r < (\rho + N + \tfrac{1}{2}) \cos\phi.$$

On considering separately the factors in (2.2) for which $1 < n < N$, $n = N$ and $n > N + 1$, as Levinson does, we get

(3.9) $$|E(w)| < \left|\frac{\Gamma(\rho + N + 1 + \delta)\, \Gamma(\rho + N + 1 - \delta - w)}{\Gamma(\rho + N + 1 - \delta)\, \Gamma(\rho + N + 1 + \delta - w)}\right| \cdot \left|\frac{\Gamma(\rho + 1 - \delta)\, \Gamma(1 - \delta)(\lambda_N - w)}{\Gamma(\rho + 1 - \delta - w)\, \Gamma(1 - \delta + w)(\rho + N - \delta - w)}\right|.$$

By Stirling's theorem the first factor in (3.9) does not exceed

(3.10) $$A_1(\rho + N - \tfrac{1}{2})^{2\delta}/|\rho + N + 1 - \delta - w|^{2\delta};$$

while the second factor does not exceed

(3.11) $$\frac{A_2}{r^{\rho+1-2\delta}} \left|\frac{(\lambda_N - w) \sin \pi(w + \delta - \rho)}{\pi(w + \delta - N - \rho)}\right|.$$

Now when $|w + \delta - N - \rho| < \tfrac{1}{2}$, $|\sin \pi(w + \delta - \rho)/\pi(w + \delta - N - \rho) < A_3$; and by (2.1), $|\lambda_N - w| < 1$. When $|w + \delta - N - \rho| > \tfrac{1}{2}$, $|\sin \pi(w + \delta - \rho)| < A_4 \exp(\pi|v|)$, and $|(\lambda_N - w)/(w + \delta - N - \rho)| < 1 + 2\delta\alpha_N/|w + \delta - N - \rho| < 2$. Thus in all cases the second factor in (3.9) does not exceed

(3.12) $$A_5 \exp(\pi|v|)/r^{\rho+1-2\delta}.$$

We prove now that

(3.13) $$bd\{|\rho + N + 1 - \delta - w| \cos\phi\} > 0.$$

For by (3.21), (3.8), when $0 < |\phi| < \pi/4$,

$$|\rho + N + 1 - \delta - w| \cos \phi > 2^{-\frac{1}{2}}(\rho + N + 1 - \delta - u) > 2^{-\frac{1}{2}}[(\rho + N + \tfrac{1}{2}) \sin^2\phi + \tfrac{1}{2} - \delta] > 2^{-\frac{1}{2}}(\tfrac{1}{2} - \delta);$$

and when $\pi/4 < |\phi| < \pi/2$,

$$\begin{aligned} |\rho + N + 1 - \delta - w| \cos \phi &> (\rho + N + 1 - \delta - u)/(\rho + N + \tfrac{1}{2}) \\ &> \sin^2\phi + (\tfrac{1}{2} - \delta)/(\rho + N + \tfrac{1}{2}) > \tfrac{1}{2}, \end{aligned}$$

by (3.8) and $r > 1$.

From (3.9), (3.10), and (3.12),

$$\begin{aligned} |E(w)| &< A_6(\rho + N - \tfrac{1}{2})^{2\delta} \exp(\pi|v|)/r^{\rho+1-2\delta}|\rho + N + 1 - \delta - w|^{2\delta}, \\ &< A_7 r^{2\delta} \exp(\pi|v|)/r^{\rho+1-2\delta}\{|\rho + N + 1 - \delta - w| \cos \phi\}^{2\delta}, \\ &< A_8 \exp(\pi|v|)/r^{\rho+1-4\delta} \end{aligned}$$

by (3.13). This proves (3.1) for $\Re(w) > 0$; and the assertion is seen to be true for $\Re(w) < 0$ by applying the same argument to $E(-w)$.

Remark on the proof of (3.4). Levinson's method may be used to show that when $\rho + N < u < \rho + 2N$,

$$(3.14) \qquad |E(u)| < \frac{A_9 N^{4\delta-\rho-1}}{(\rho + 2N + 1 - u)^{2\delta}} \prod_{N}^{2N} \left| \frac{\rho + n - \delta + \alpha_n - u - 2i}{\rho + n - \delta - u - 2i} \right|^{2\delta},$$

$$(3.15) \qquad \int_{\rho+N}^{\rho+2N} |E(u)|^q du < A_{10}/N^{q(\rho+1)-1-2q\delta},$$

provided that

$$(3.16) \qquad 0 < 2q\delta < 1.$$

If in addition

$$(3.17) \qquad q(\rho + 1 - 2\delta) > 1,$$

it follows from (3.15) that $E(u) \in L(0, \infty)$. The final conclusion follows by considering $E(-u)$ in the same way.

It is evident that there always exists a number q, $1 < q < 2$, satisfying (3.16) and (3.17), for example $q = (1 - \Delta)^{-1}$, where $2\delta < \Delta < 1 - 2\delta$.

Proof of (3.5). Since

$$\begin{aligned} \left|\left(1 - \frac{iv}{\rho + n + \delta}\right)\left(1 + \frac{iv}{n + \delta}\right)\right| &< \left|\left(1 - \frac{iv}{\lambda_n}\right)\left(1 + \frac{iv}{\mu_n}\right)\right| \\ &< \left|\left(1 - \frac{iv}{\rho + n - \delta}\right)\left(1 + \frac{iv}{n - \delta}\right)\right|, \end{aligned}$$

it follows that

$$\begin{aligned} \left|\frac{\Gamma(\rho + 1 + \delta)\, \Gamma(1 + \delta)}{\Gamma(\rho + 1 + \delta - iv)\, \Gamma(1 + \delta + iv)}\right| &< |E(iv)| \\ &< \left|\frac{\Gamma(\rho + 1 - \delta)\, \Gamma(1 - \delta)}{\Gamma(\rho + 1 - \delta - iv)\, \Gamma(1 - \delta + iv)}\right|, \end{aligned}$$

and (3.5) then follows from a classical property (9, p. 259) of the Γ-function.

Proof of (3.6). We give details for the case $v > 0$. Writing

$$\phi_n = \operatorname{amp}\{(1 - iv\theta/\lambda_n)(1 + iv\theta/\mu_n)/(1 - iv/\lambda_n)(1 + iv/\mu_n)\},$$

and using the inequalities for λ_n, μ_n and ρ in (2.1), we have

$$\begin{aligned}\phi_n &= \arctan\left(\frac{v(1-\theta)\,\lambda_n}{v^2\theta + \lambda_n^2}\right) - \arctan\left(\frac{v(1-\theta)\,\mu_n}{v^2\theta + \mu_n^2}\right),\\ &< \arctan\left(\frac{v(1-\theta)(\rho + n + \delta)}{v^2\theta + (\rho + n - \delta)^2}\right) - \arctan\left(\frac{v(1-\theta)(n-\delta)}{v^2\theta + (n+\delta)^2}\right),\\ &< \arctan\left(\frac{v(1-\theta)(n + 1 - \delta)}{v^2\theta + (n - \delta)^2}\right) - \arctan\left(\frac{v(1-\theta)(n-\delta)}{v^2\theta + (n+\delta)^2}\right),\\ &= \arctan\frac{v(1-\theta)\{v^2\theta + n^2(1+4\delta) + 2n\delta(1-2\delta) + \delta^2\}}{[v^2\theta + (n-\delta)^2][v^2\theta + (n+\delta)^2] + v^2(1-\theta)^2(n+1-\delta)(n-\delta)}.\end{aligned}$$

On observing that

$$0 < v^2\theta + n^2(1 + 4\delta) + 2n\delta(1 - 2\delta) + \delta^2 < 2[v^2\theta + (n + \delta)^2],$$

we see that $\phi_n < \arctan\{2v(1 - \theta)/[v^2\theta + (n - \delta)^2]\}$. A similar argument applied to $-\phi_n$ gives

$$-\phi_n < \arctan\{3v(1 - \theta)/[v^2\theta + (n - \delta)^2]\};$$

and thus

$$|\phi_n| < \arctan\{3v(1 - \theta)/[v^2\theta + (n - \delta)^2]\}.$$

It then follows easily that $|\operatorname{amp}[E(iv\theta)/E(iv)]| < A(1 - \theta)$, the constant being independent of v.

Proof of (3.7). Let Λ be the region consisting of the w-plane from which the points $v = 0$, $|u| > 1 - \delta$ have been removed. Then the series

$$\sum_1^\infty \left(\frac{w}{\lambda_n - w} - \frac{\theta w}{\lambda_n - \theta w}\right), \quad \sum_1^\infty \left(\frac{w}{\mu_n - w} - \frac{\theta w}{\mu_n - \theta w}\right)$$

converge absolutely and uniformly in any compact subset of Λ, and

$$\begin{aligned}\frac{d}{dv}\log\left[\frac{E(iv\theta)}{E(iv)}\right] &= i\sum_1^\infty \left(\frac{\lambda_n + iv}{\lambda_n^2 + v^2} - \frac{\theta(\lambda_n + iv\theta)}{\lambda_n^2 + v^2\theta^2}\right)\\ &\quad - i\sum_1^\infty \left(\frac{\mu_n - iv}{\mu_n^2 + v^2} - \frac{\theta(\mu_n - iv\theta)}{\mu_n^2 + v^2\theta^2}\right),\end{aligned}$$

$$\Re\left\{\frac{d}{dv}\log\frac{E(iv\theta)}{E(iv)}\right\} = -\sum_1^\infty \left[\frac{v\lambda_n^2(1 - \theta^2)}{(\lambda_n^2 + v^2)(\lambda_n^2 + v^2\theta^2)} + \frac{v\mu_n^2(1 - \theta^2)}{(\mu_n^2 + v^2)(\mu_n^2 + v^2\theta^2)}\right] < 0.$$

It follows that $|E(iv\theta)/E(iv)|$ is a decreasing function of $|v|$.

4. Representation of the operator. Let p be the index conjugate to q, so that $p > 2$. By (3.4), the function

$$k(y) = \operatorname*{lim^{(p)}}_{R\to\infty} (2\pi)^{-\frac{1}{2}} \int_{-R}^{R} E(u)\exp(-iuy)\,du$$

exists and belongs to $L^p(-\infty, \infty)$. On considering the contour integral $\int E(w) \exp(-iyw)\, dw$ taken round the boundary of the semi-circular disc $|w| < R$, $v > 0$, we see that the part of the integral taken round the arc of the semi-circle is $O(R^{-\beta})$, $(R \to \infty)$ when $|y| > \pi$, where $\beta = \rho + 1 - 4\delta > 0$ in (3.1). Thus

$$k(y) = 0 \qquad p.p. \text{ in } |y| > \pi. \tag{4.1}$$

Since $E(u)$ is continuous

$$\begin{aligned} E(u) &= \lim_{R\to\infty} (2\pi)^{-\frac{1}{2}} \int_{-R}^{R} k(y)[1 - |y|/R] \exp(iuy)\, dy, \\ &= (2\pi)^{-\frac{1}{2}} \int_{-\pi}^{\pi} k(y) \exp(iuy)\, dy, \end{aligned} \tag{4.2}$$

and $k(y) \in L^2(-\pi, \pi)$ by (4.1) and $p > 2$.

It is easily seen that for complex w, the integral

$$(2\pi)^{-\frac{1}{2}} \int_{-\pi}^{\pi} k(y) \exp(iyw)\, dy$$

defines an entire function, and by (4.2) we may write

$$E(w) = (2\pi)^{-\frac{1}{2}} \int_{-\pi}^{\pi} k(y) \exp(iyw)\, dy. \tag{4.3}$$

In proving the inversion theorem we shall make use of the function $E(\theta w)$, $0 < \theta < 1$, for which we prove

$$|E(\theta r_n)| < A(1 - \theta) r_n, \tag{4.4}$$

where r_n stands for λ_n or $-\mu_n$ and the constant is independent of n.

Let $m(y, \alpha) = \exp(-iy\alpha\theta) - \exp(-iy\alpha)$. Then for $|y| < \pi$, and $0 < \theta < 1$,

$$|m(y, \alpha)| = |\alpha| \left| \int_{\theta y}^{y} \exp(-it\alpha)\, dt \right| < \pi|\alpha|(1 - \theta);$$

$$|E(-\theta\mu_k) - E(-\mu_k)| = (2\pi)^{-\frac{1}{2}} \left| \int_{-\pi}^{\pi} k(y)\, m(y, \mu_k)\, dy \right| < A(1 - \theta)\, \mu_k,$$

by the Schwarz inequality.

With the usual interpretation of $\exp(aD)$ as a shift operator, we have formally

$$\begin{aligned} DE(D).f(x) &= \lim_{\theta\to 1} DE(\theta D).f(x), \\ &= \lim_{\theta\to 1} (2\pi)^{-\frac{1}{2}} \int_{-\pi}^{\pi} k(y) \exp(iy\theta D)\, dy.f'(x), \\ &= \lim_{\theta\to 1} (2\pi)^{-\frac{1}{2}} \int_{-\pi}^{\pi} k(y) f'(x + iy\theta)\, dy. \end{aligned}$$

We therefore *define* the operation $DE(D).f(x)$ by

$$DE(D).f(x) = \lim_{\theta\to 1} (2\pi)^{-\frac{1}{2}} \int_{-\pi}^{\pi} k(y) f'(x + iy\theta)\, dy. \tag{4.5}$$

5. Properties of the nucleus. Denoting the strip $|y| < \pi$ of the z-plane by B, we prove the following propositions:

(5.1) *the integral* (2.3) *defining* $G(z)$ *converges absolutely when* $z \in B$, *converges uniformly when* z *belongs to a compact subset of* B, *and therefore defines a function analytic in* B;

$$G(z) = \begin{cases} 1 + O[\exp(-x\mu_1)], & (x \to \infty \text{ in } B), \\ O[\exp(x\lambda_1)], & (x \to -\infty \text{ in } B); \end{cases} \tag{5.2}$$

$$G'(z) = \begin{cases} O[\exp(-x\mu_1)], & (x \to \infty \text{ in } B), \\ O[\exp(x\lambda_1)], & (x \to -\infty \text{ in } B); \end{cases} \tag{5.3}$$

(5.4) *when* $z_0 \in B$, *there is a constant* R_0 *such that the integrals*

$$\int_{-\infty}^{-R_0} \left| \frac{d}{dt} \frac{G(z-t)}{G(z_0-t)} \right| dt, \quad \int_{R_0}^{\infty} \left| \frac{d}{dt} \frac{G(z-t)}{G(z_0-t)} \right| dt$$

converge uniformly when z *belongs to any compact subset of* B.

Proof of (5.1). By (2.3), (3.2),

$$\left| \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{\exp(zw)\, dw}{wE(w)} \right| < A_1 \exp(cx) \int_{-\infty}^{\infty} \exp[-yv - (\pi - \epsilon)|v|]\, dv$$
$$< A_2 \exp(cx),$$

since $|y| < \pi$ in B. This inequality is sufficient to establish the assertions of (5.1).

Proof of (5.2), (5.3). On account of the classical properties of Dirichlet series it is sufficient to show that

$$G(z) = \begin{cases} 1 - \sum_1^\infty \exp(-z\mu_n)/\mu_n E'(-\mu_n), & (x > 0, |y| < \pi), \\ - \sum_1^\infty \exp(z\lambda_n)/\lambda_n E'(\lambda_n), & (x < 0, |y| < \pi), \end{cases} \tag{5.5}$$

the Dirichlet series converging absolutely in the indicated regions.

Details are given for the case $x < 0$. Designating the points $c - iR$, $c + R \cot \beta - iR$, $c + R \cot \beta + iR$ and $c + iR$ by A, B, C and D respectively, where $0 < \beta < \frac{1}{2}\pi$, let L be the contour formed by the linear segments AB and CD and the circular arc $|w - c| = R$ joining to B to C. Consider

$$I = \int_L \exp(zw)\, dw/wE(w).$$

By using (3.2) and $x < 0$, $0 < \beta < \frac{1}{2}\pi$, $|y| < \pi$, and estimating the integrals along AB, BC and CD separately, we see that $I = o(1)$ as $R \to \infty$. The second equation in (5.5) then follows from the definition of $G(z)$ and the calculus of residues.

Proof of (5.4). When $x \neq 0$, $G(z)$ is represented by the absolutely convergent Dirichlet series (5.5), and it is well known that functions so defined can have but a finite number of zeros. Since z_0 is given, and $X_1 < x < X_2$

in the compact subset of B, we may choose R_0 so that $G(z_0 - t)$ does not vanish for $|t| > R$. It then follows easily that

$$\left|\frac{d}{dt}\frac{G(z-t)}{G(z_0-t)}\right| = \begin{cases} O[\exp(t\mu_1)] & (t \to -\infty), \\ O\{\exp[-t(\lambda_2 - \lambda_1)]\} & (t \to \infty), \end{cases}$$

where the constants are independent of z. These estimates are sufficient for the proof.

6. Properties of the transform. The following theorem gives properties of the functions $f(x)$ and $\phi(t)$ in (1.1) which will be used later.

THEOREM I. *Let $\phi(t) \in L(0, R)$ for any R and be such that the integral* (1.1) *converges for at least one z in B, and let $\Phi(t) = \int_0^t \phi(u)du$: then*
(6.1) *the integral* (1.1) *converges for all z in B, and defines a function analytic in B;*

$$\Phi(t) = \begin{cases} o(\exp t\lambda_1) & (t \to \infty), \\ o(\exp - t\Delta) & (t \to -\infty), \end{cases} \tag{6.2}$$

for any positive Δ.

Proof of (6.1). On account of (5.4), the method of Widder-Hirschmann may be used (**4**, pp. 691–692).

Proof of (6.2). It follows from the representation (5.5) of $G(z)$ for $x \neq 0$ that $G(z)$ and $G'(z)$ have at most a finite number of zeros. Let A be a positive number such that neither $G(z)$ nor $G'(z)$ vanishes for $|x| > A$. To prove the first assertion, define $\psi(t) = \int_0^t G(-A-u)\, d\Phi(u)$. Then by hypothesis, $\psi(\infty)$ is finite, and

$$\begin{aligned} \Phi(t)\exp(-t\lambda_1) &= \exp(-t\lambda_1) \int_0^t d\psi(u)/G(-A-u), \\ &= \exp(-t\lambda_1)\left[\frac{\psi(t)}{G(-A-t)} - \int_0^t \frac{\psi(u)\, G'(-A-u)\, du}{G^2(-A-u)}\right], \\ &= o(1) \end{aligned}$$

as $t \to \infty$ by l'Hopital's rule, and (5.5).

To prove the second assertion, write $\psi(t) = \int_0^t G(A-u)\, d\Phi(u)$; and in the same way, $\Phi(t)\exp(t\Delta) = o(1)$, $(t \to -\infty)$.

It is convenient at this point to establish some properties of the function

$$K(x, \theta) = \theta(2\pi)^{-1} \int_{-\pi}^{\pi} k(y)\, G'(x + iy\theta)\, dy. \tag{6.3}$$

These properties are:

$$K(x, \theta) = \begin{cases} (1-\theta)\, O[\exp(-x\mu_1)], & (x \to \infty), \\ (1-\theta)\, O[\exp(x\lambda_1)], & (x \to -\infty), \end{cases} \tag{6.4}$$

with similar estimates for $K'(x, \theta)$;

(6.5) $K(x, \theta) = O[(1-\theta)^{-1}]$ *uniformly in x as $\theta \to 1$;*

(6.6) *when x is positive,* $\lim_{\theta \to 1} \int_0^x K(t, \theta)\, dt = \frac{1}{2} = \lim_{\theta \to 1} \int_{-x}^0 K(t, \theta)\, dt.$

Proof of (6.4). Details are given for the case $x > 0$.

$$|K(x,\theta)| = \left|\theta(2\pi)^{-\frac{1}{2}}\int_{-\pi}^{\pi} k(y)\,dy\,(2\pi i)^{-1}\int_{c-i\infty}^{c+i\infty}\exp[w(x+iy\theta)]\,dw/E(w)\right|$$

$$= \left|(2\pi i)^{-1}\int_{c-i\infty}^{c+i\infty}E(\theta w)\exp(xw)\,dw/E(w)\right|, \qquad \textit{by } (4.3);$$

$$= \left|\sum_{1}^{\infty}E(-\theta\mu_n)\exp(-x\mu_n)/E'(-\mu_n)\right|,$$

$$< A_1(1-\theta)\sum_{1}^{\infty}\mu_n\exp(-x\mu_n)/|E'(-\mu_n)|, \qquad \textit{by } (4.4);$$

and as $x > 0$ and (3.3) guarantee the convergence of this series, our assertion is proved.

Proof of (6.5). By (4.3), (6.3) and Cauchy's theorem,

$$K(x,\theta) = \theta(2\pi)^{-\frac{1}{2}}\int_{-\pi}^{\pi}k(y)\,dy\,(2\pi i)^{-1}\int_{-i\infty}^{i\infty}\exp[w(x+iy\theta)]dw/E(w),$$

$$= \theta(2\pi)^{-1}\int_{-\infty}^{\infty}\exp(ixv)\,E(iv\theta)\,dv/E(iv).$$

Using the fact that for $0 < \theta < 1$,

$$|(1 - iv\theta/\lambda_n)/(1 - iv\lambda_n)| \text{ and } |(1 + iv\theta/\mu_n)/(1 + iv/\mu_n)|$$

are less than unity, we have

$$\left|\frac{1 - iv\theta/(\rho+n-\delta)}{1 - iv/(\rho+n-\delta)}\right| < \left|\frac{1 - iv\theta/\lambda_n}{1 - iv/\lambda_n}\right| < \left|\frac{1 - iv\theta/(\rho+n+\delta)}{1 - iv/(\rho+n+\delta)}\right|,$$

with similar inequalities involving μ_n. Hence

$$\left|\frac{\Gamma(\rho+1-\delta-iv)\,\Gamma(1-\delta+iv)}{\Gamma(\rho+1-\delta-iv\theta)\,\Gamma(1-\delta+iv\theta)}\right| < \left|\frac{E(iv\theta)}{E(iv)}\right|$$

$$< \left|\frac{\Gamma(\rho+1+\delta-iv)\,\Gamma(1+\delta+iv)}{\Gamma(\rho+1+\delta-iv\theta)\,\Gamma(1+\delta+iv\theta)}\right|,$$

and by (9, p. 259),

$$(6.7) \qquad |E(iv\theta)/E(iv)| \sim A\exp[-\pi|v|(1-\theta)] \qquad (|v| \to \infty).$$

Since $0 < \theta < 1$, this is sufficient to prove our result.

Proof of (6.6). Write $I = \int_0^x K(t,\theta)\,dt$, where $x \neq 0$. Then

$$I = \theta(2\pi)^{-\frac{1}{2}}\int_0^x dt\int_{-\pi}^{\pi}k(y)\,G'(t+iy\theta)\,dy,$$

$$= \theta(2\pi)^{-\frac{1}{2}}\int_{-\pi}^{\pi}k(y)[G(x+iy\theta) - G(iy\theta)]\,dy,$$

the interchange of the integrations being justified, since $k(y) \in L^2(-\pi,\pi)$ and

$$\begin{aligned}
|G'(t+iy\theta)| &= \left|(2\pi i)^{-1}\int_{c-i\infty}^{c+i\infty}\exp[w(t+iy\theta)]\,dw/E(w)\right|,\\
&= \left|(2\pi)^{-1}\int_{-\infty}^{\infty}\exp(-vy\theta+ivt)\,dv/E(iv)\right|,\\
&= \int_{-\infty}^{\infty}\exp[-\pi|v|(1-\theta)]\,O[|v|^{\rho+1+2\delta}]\,dv, \textit{ by } (3.5),\\
&= O[(1-\theta)^{-\rho-2-2\delta}].
\end{aligned}$$

Thus

$$I = \theta(2\pi)^{-\frac{1}{2}}\int_{-\pi}^{\pi}k(y)dy(2\pi i)^{-1}\int_{c-i\infty}^{c+i\infty}\{\exp[w(x+iy\theta)]-\exp(iy\theta w)\}dw/wE(w).$$

Again by (4.3) and the absolute convergence of the inner integral for $|y| < \pi$, we may interchange the integrations, and get

$$I = \frac{\theta}{2\pi}\int_{-\infty}^{\infty}\frac{[\exp(ixv)-1]\,E(iv\theta)}{ivE(iv)}\,dv, \tag{6.8}$$

the application of Cauchy's theorem being justified by the analyticity of $[\exp(xw)-1]/w$ at the origin.

We observe next that $E(\theta w)/E(w)$ is real when w is real. Hence $I = P - Q$, where

$$\begin{aligned}
Q &= \frac{\theta}{2\pi}\int_0^{\infty}\frac{1-\cos xv}{v}\,\mathfrak{I}\left[\frac{E(iv\theta)}{E(iv)}\right]dv,\\
P &= \frac{\theta}{2\pi}\int_0^{\infty}\frac{\sin xv}{v}\,\mathfrak{R}\left[\frac{E(iv\theta)}{E(iv)}\right]dv.
\end{aligned}$$

It is then sufficient to show that

$$Q \to 0, \qquad \theta \to 1, \tag{6.9}$$

$$P \to \tfrac{1}{2}, \qquad \theta \to 1. \tag{6.10}$$

To prove (6.9), it is sufficient to consider

$$Q_1 = \int_1^{\infty}\frac{1-\cos xv}{v}\,\mathfrak{I}\left[\frac{E(iv\theta)}{E(iv)}\right]dv.$$

We then have

$$\begin{aligned}
|Q_1| &= \left|\int_1^{\infty}\frac{1-\cos xv}{v}\left|\frac{E(iv\theta)}{E(iv)}\right|\sin\operatorname{amp}\left[\frac{E(iv\theta)}{E(iv)}\right]dv\right|,\\
&< A_1(1-\theta)\int_1^{\infty}v^{-\frac{1}{2}}\exp[-\pi v(1-\theta)]\,dv, \qquad \textit{by } (3.6) \textit{ and } (6.7),\\
&= A_1(1-\theta)\int_{\pi(1-\theta)}^{\infty}t^{-\frac{1}{2}}\exp(-t)\,dt < A_2(1-\theta)^{\frac{1}{2}}\Gamma(\tfrac{1}{2}).
\end{aligned}$$

Thus Q_1, and consequently Q tends to zero as θ tends to unity.

To prove (6.10), we observe that on account of (3.7),

$$\int_0^{\infty}\frac{\sin xv}{v}\left|\frac{E(iv\theta)}{E(iv)}\right|dv$$

converges uniformly in $\frac{1}{2} < \theta < 1$; and from (3.6), that $\mathfrak{R}[E(iv\theta)/E(iv)]$ is

positive when θ is close to unity. It is therefore sufficient to prove that $C(\theta) \to 0$ $(\theta \to 1)$, where

$$C(\theta) = \int_0^\infty \frac{\sin xv}{v}\left\{\left|\frac{E(iv\theta)}{E(iv)}\right| - \Re\left[\frac{E(iv\theta)}{E(iv)}\right]\right\} dv.$$

But

$$|C(\theta)| < \int_0^\infty \left|\frac{E(iv\theta)}{E(iv)}\right|\left\{1 - \cos \operatorname{amp}\left[\frac{E(iv\theta)}{E(iv)}\right]\right\} dv,$$

$$< A_3(1-\theta)^2 \int_0^\infty \exp[-\pi v(1-\theta)]\, dv,$$

$$= O[(1-\theta)], \qquad \text{by (3.6).}$$

7. The inversion theorems. The main result is

THEOREM II. *Let $\phi(t) \in L(0, R)$ for any R and be such that the integral* (1.1) *converges for at least one z in the strip B: then if $f(z)$ is defined by* (1.1) *and $DE(D).f(x)$ by* (4.5),

$$DE(D).f(x) = \tfrac{1}{2}[\phi(x+) + \phi(x-)],$$

whenever the right-hand side has a meaning.

For

$$DE(D).f(x) = \lim_{\theta\to1} (2\pi)^{-\frac{1}{2}} \int_{-\pi}^{\pi} k(y)\, dy \int_{-\infty}^{\infty} G'(x - t + iy\theta)\, \phi(t)\, dt,$$

$$= \lim_{\theta\to1} \int_{-\infty}^{\infty} \phi(t)\, dt\, (2\pi)^{-\frac{1}{2}} \int_{-\pi}^{\pi} k(y)\, G'(x - t + iy\theta)\, dy,$$

$$= \lim_{\theta\to1} \int_{-\infty}^{\infty} K(x - t, \theta)\, \phi(t)\, dt,$$

the interchange of the integrations being justified by the uniform convergence of (1.1) in any compact subset of B, and the fact that $k(y) \in L^2(-\pi, \pi)$. It is sufficient to prove

$$\int_{-\infty}^{x} K(x - t, \theta)\, \phi(t)\, dt \to \tfrac{1}{2}\phi(x-), \tag{7.1}$$

$$\int_{x}^{\infty} K(x - t, \theta)\, \phi(t)\, dt \to \tfrac{1}{2}\phi(x+), \tag{7.2}$$

as $\theta \to 1$.

We give details for (7.1). Let $T > 0$, and write

$$\int_{x-T}^{x} K(x-t, \theta)[\phi(t) - \phi(x-)]dt = \int_0^T K(t, \theta)[\phi(x-t) - \phi(x-)]dt = J(0, T).$$

Then

$$|J[0, \pi(1-\theta)]| < \int_0^{\pi(1-\theta)} |\phi(x-t) - \phi(x-)||K(t,\theta)|dt$$

$$< A(1-\theta)^{-1} \int_0^{\pi(1-\theta)} |\phi(x-t) - \phi(x-)|dt$$

$$= o(1), \tag{7.3}$$

as $\theta \to 1$, by (6.5). Next by (6.4),

$$|J[\pi(1-\theta), T]| < A(1-\theta) \int_{\pi(1-\theta)}^{T} |\phi(x-t) - \phi(x-)| \exp(-t\mu_1)\,dt \tag{7.4}$$
$$= O(1-\theta).$$

Thus by (6.6), (7.3) and (7.4)

$$\lim_{\theta \to 1} \int_{x-T}^{x} K(x-t, \theta)\, \phi(t)\, dt = \tfrac{1}{2}\phi(x-).$$

It remains to prove that $\int_{-\infty}^{x-T} K(x-t,\theta)\,\phi(t)\,dt \to 0$ as $\theta \to 1$. As this integral need not converge absolutely, we write it as

$$[K'(x-t,\theta)\,\Phi(t)]_{-\infty}^{T} + \int_{-\infty}^{x-T} K'(x-t,\theta)\,\Phi(t)\,dt.$$

By (6.2) and (6.4) the integrated term $= o(1)$; and for the same reason

$$\left| \int_{-\infty}^{x-T} K'(x-t,\theta)\,\Phi(t)\,dt \right| = \left| \int_{T}^{\infty} K'(t,\theta)\,\Phi(x-t)\,dt \right| = O(1-\theta).$$

Since (7.2) may be proved in the same way, the theorem is complete.

The proof of the following theorem is similar:

THEOREM III. *Let $f(z) = \int_{-\infty}^{\infty} G(z-t)\,d\alpha(t)$, where $\alpha(t)$ is a normalized function of bounded variation in any finite interval: then if this integral converges for any z in B, it converges for all such z, converges uniformly in any compact subset of B, and defines a function analytic in B. Also*

$$\lim_{\theta \to 1} (2\pi)^{-1} \int_{x_1}^{x_2} dx \int_{-\pi}^{\pi} k(y)\, f'(x+iy\theta)\, dy = \alpha(x_2) - \alpha(x_1).$$

8. Remarks. In the proof of (5.4) we have used the fact that from its representation (5.5) as a Dirichlet series, the nucleus $G(z)$ has but a finite number of zeros. Hirschmann and Widder (**8**, p. 159) have shown that a more general nucleus has no zeros on the real axis, and it is certainly true that $G(iy) \neq 0$ for $|y| < \pi$. The proof that $G(z)$ does not vanish in B seems to be connected with properties of functions defined by Dirichlet series with coefficients of alternating sign, and will be dealt with elsewhere.

REFERENCES

1. N. Wiener, *The operational calculus*, Math. Ann., *95* (1925–26), 557–584.
2. D. V. Widder, *Inversion formulas for convolution transforms*, Duke Math. J., *14* (1947), 217–249.
3. ———, *The Stieltjes transform*, Trans. Amer. Math. Soc., *43* (1938), 7–60.
4. I. I. Hirschmann, Jr. and D. V. Widder, *Generalized inversion formulas for convolution transforms*, Duke Math. J., *15* (1948), 659–696.
5. D. B. Sumner, *An inversion formula for the generalized Stieltjes transform*, Bull. Amer. Math. Soc., *55* (1949), 174–183.

6. I. I. Hirschmann, Jr. and D. V. Widder, *Convolution transforms with complex kernels*, Pacific J. Math., *1* (1951), 211–225.
7. N. Levinson, *Gap and density theorems* (Amer. Math. Soc. Coll., vol. XXVI, 1940).
8. I. I. Hirschmann, Jr. and D. V. Widder, *Inversion of a class of convolution transforms*, Trans. Amer. Math. Soc., *66* (1949), 135–201.
9. E. C. Titchmarsh, *The theory of functions* (1st. ed., Oxford, 1932).
10. G. H. Hardy and N. Levinson, *Inequalities satisfied by a certain definite integral*, Bull. Amer. Math. Soc., *43* (1937), 709–716.

Hamilton College, McMaster University, and
Research Institute of the Canadian Mathematical Congress.

ON AN INVERSION FORMULA FOR THE LAPLACE TRANSFORMATION

P. G. ROONEY

1. Introduction. The literature of the Laplace transformation contains many examples of inversion operators. Particular attention has been given in this literature to the so-called "real" inversion operators, that is, those operators which make use of values of the generating function arising only from real values of the independent variable.

It is one of these "real" inversion operators which we shall study here, one which has already been studied in part. If

$$\text{I} \qquad f(s) = \int_0^\infty e^{-st}\phi(t)\,dt$$

then the inversion operator is defined by the formula

$$\text{II} \qquad {}^{\nu}L_{k,t}[f(s)] = \frac{k^{3/2}e^{2k}}{t\pi^{\frac{1}{2}}}\int_0^\infty x^{\frac{1}{2}\nu}J_\nu(2kx^{\frac{1}{2}})\,f\left(\frac{k(x+1)}{t}\right)dx$$

and we shall show that, under certain conditions,

$$\lim_{k\to\infty} {}^{\nu}L_{k,t}[f(s)] = \phi(t).$$

The operator was originally given by Erdélyi **(1)**. However the resulting inversion and representation theories were not studied there. A special case of the operator, namely $\nu = -\frac{1}{2}$, was studied by the author in **(2)**.

We shall derive our results here partially from those of **(2)** by means of a relation, also derived here, between the operators arising from different values of ν.

In §2 we find conditions for the existence of the operator, and derive a number of its properties, including the one mentioned in the previous paragraph.

Section 3 contains the inversion theory, and in § 4 we find necessary and sufficient conditions that a function be equal almost everywhere to a Laplace transform of a function of the form $t^{\alpha}\phi(t)$, $\phi \in L_p(0, \infty)$, $1 < p \leqslant \infty$, $\alpha > -(p-1)/p$.

The notations introduced by formulas I and II will be taken as standard throughout this paper, as well as

$$\text{III} \qquad f_\mu(s) = \int_0^\infty e^{-st}t^{-\mu}\phi(t)\,dt,$$

$$\text{IV} \qquad \bar{f}(s) = \int_0^\infty e^{-st}|\phi(t)|\,dt.$$

Received December 11, 1953.

2. Existence and properties of the operator. The following theorem gives sufficient conditions for the operator to exist for a given $\phi(t)$.

THEOREM 2.1. *If*

(1) $$e^{-\gamma t}\phi(t) \in L(0, \infty), \quad \gamma > 0,$$
(2) $$t^{-(\frac{1}{2}\nu+3/4)}\phi(t) \in L(0, \delta), \text{ for some } \delta > 0,$$
(3) $$\nu > -1,$$

then for each $t > 0$, and all $k > \gamma t$, ${}^{\nu}L_{k,t}[\bar{f}(s)]$ and ${}^{\nu}L_{k,t}[f(s)]$ exist.

Proof. Clearly the existence of ${}^{\nu}L_{k,t}[\bar{f}(s)]$ implies that of ${}^{\nu}L_{k,t}[f(s)]$, since $|f(s)| < \bar{f}(s)$. Let $k > \gamma t$, and $t > 0$. Now by (**3**; §7.21)

$$J_\nu(x) \sim \left(\frac{2}{\pi x}\right)^{\frac{1}{2}} \cos(x - \tfrac{1}{2}\nu\pi - \tfrac{1}{4}\pi) \text{ as } x \to \infty,$$

and hence $R > 0$ and M exist such that $|J_\nu(x)| < Mx^{-\frac{1}{2}}$ for $x > R$. Thus

$$\begin{aligned}&\int_{R/2k}^{\infty} e^{-kuy^2/t} y^{\nu+1} |J_\nu(2ky)|\, dy \\ &\qquad < (2k)^{-\frac{1}{2}} M \int_0^\infty e^{-kuy^2/t} y^{\nu+\frac{1}{2}} dy = (2/k)^{\frac{1}{2}} M\Gamma(\tfrac{1}{2}\nu + \tfrac{3}{4})(ku/t)^{-(\frac{1}{2}\nu+3/4)}.\end{aligned}$$

Further, since $\nu > -1$, the integral,

$$\int_0^{R/2k} e^{-kuy^2/t} y^{\nu+1} |J_\nu(2ky)|\, dy$$

is uniformly bounded in u for $u > 0$. Hence,

$$\begin{aligned}&\int_0^\infty e^{-ku/t}|\phi(u)|\, du \int_0^\infty e^{-kuy^2/t} y^{\nu+1}|J_\nu(2ky)|\, dy \\ &\quad = \int_0^\infty e^{-ku/t}|\phi(u)|\, du \left\{ \int_0^{R/2k} + \int_{R/2k}^{\infty} \right\} e^{-kuy^2/t} y^{\nu+1} |J_\nu(2ky)|\, dy \\ &\quad < M_1 \int_0^\infty e^{-ku/t}|\phi(u)|\, du + M_2 \int_0^\infty e^{-ku/t} u^{-(\frac{1}{2}\nu+3/4)} |\phi(u)|\, du < \infty,\end{aligned}$$

by (1) and (2), and where

$$M_1 = \int_0^{R/2k} y^{\nu+1}|J_\nu(2ky)|\, dy, \text{ and } M_2 = M\Gamma(\tfrac{1}{2}\nu + \tfrac{3}{4})(2/k)^{\frac{1}{2}}(ku/t)^{-(\frac{1}{2}\nu+3/4)}.$$

Hence, by Fubini's theorem we have that

$$\begin{aligned}&\frac{2k^{3/2}e^{2k}}{t\pi^{\frac{1}{2}}} \int_0^\infty e^{-ku/t}|\phi(u)|\, du \int_0^\infty e^{-kuy^2/t} y^{\nu+1} J_\nu(2ky)\, dy \\ &\quad = \frac{2k^{3/2}e^{2k}}{t\pi^{\frac{1}{2}}} \int_0^\infty y^{\nu+1} J_\nu(2ky)\, dy \int_0^\infty e^{-k(y^2+1)u/t}|\phi(u)|\, du \\ &\quad = \frac{k^{3/2}e^{2k}}{t\pi^{\frac{1}{2}}} \int_0^\infty x^{\frac{1}{2}\nu} J_\nu(2kx^{\frac{1}{2}}) \int_0^\infty e^{-k(x+1)u/t}|\phi(u)|\, du \qquad (\text{where } x = y^2) \\ &\quad = \frac{k^{3/2}e^{2k}}{t\pi^{\frac{1}{2}}} \int_0^\infty x^{\frac{1}{2}\nu} J_\nu(2kx^{\frac{1}{2}})\, \bar{f}\left(\frac{k(x+1)}{t}\right) dx = {}^{\nu}L_{k,t}[\bar{f}(s)]\end{aligned}$$

exists for each $t > 0$ and all $k > \gamma t$.

COROLLARY. *If*

(1) $$e^{-\gamma t}\phi(t) \in L(0, \infty), \ \gamma > 0,$$

(2) $$\nu > -1,$$

then for each $t > 0$ *and all* $k > \gamma t$, *and each* $\epsilon > 0$, ${}^{\nu}L_{k,t}[e^{-\epsilon s}f(s)]$ *exists.*

Proof. $e^{-\epsilon s}f(s) = \int_0^\infty e^{-st}\phi_\epsilon(t)\,dt$, where $\phi_\epsilon(t) = 0$, for $0 < t < \epsilon$, and $\phi_\epsilon(t) = \phi(t - \epsilon)$ for $t > \epsilon$. Thus the conditions of Theorem 2.1 are clearly fulfilled relative to $\phi_\epsilon(t)$, and ${}^{\nu}L_{k,t}[e^{-\epsilon s}f(s)]$ exists.

The next theorem relates the operators arising from different values of ν.

THEOREM 2.2. *If*

(1) $$e^{-\gamma t}\phi(t) \in L(0, \infty), \quad \gamma > 0,$$

(2) $$t^{-(\frac{1}{2}\lambda+\mu+3/4)}\,\phi(t) \in L(0, \delta) \text{ for some } \delta > 0,$$

(3) $$\lambda > -1, \quad \mu > 0,$$

then $f_\mu(s)$ *exists for* $s > \gamma$,

$${}^{\lambda+\mu}L_{k,t}[f(s)] \text{ and } {}^{\lambda}L_{k,t}[f_\mu(s)]$$

exists for each $t > 0$ *and all* $k > \gamma t$, *and for all such* k *and* t,

$${}^{\lambda+\mu}L_{k,t}[f(s)] = t^{\mu}\,{}^{\lambda}L_{k,t}[f_\mu(s)].$$

Proof. The existence of $f_\mu(s)$ is clear. The existence of

$${}^{\lambda+\mu}L_{k,t}[f(s)] \text{ and of } {}^{\lambda}L_{k,t}[f_\mu(s)]$$

follows from Theorem 2.1.

Now by (3; §12.11), for $\lambda > -1, \mu > 0$,

$$J_{\lambda+\mu}(z) = \frac{z^\mu}{2^{\mu-1}\Gamma(\mu)}\int_0^{\frac{1}{2}\pi} J_\lambda(z\sin\theta)\sin^{\lambda+1}\theta\cos^{2\mu-1}\theta\,d\theta.$$

Hence, setting $z = 2kx^{\frac{1}{2}}$, and $z\sin\theta = 2ku^{\frac{1}{2}}$, we have

$$x^{\frac{1}{2}(\lambda+\mu)}J_{\lambda+\mu}(2kx^{\frac{1}{2}}) = \frac{k^\mu}{\Gamma(\mu)}\int_0^x u^{\frac{1}{2}\lambda}J_\lambda(2ku^{\frac{1}{2}})(x-u)^{\mu-1}du.$$

Further, for $\mu > 0$, $u \geqslant 0$, and $k \geqslant \gamma t > 0$,

$$\begin{aligned}\int_u^\infty (x-u)^{\mu-1}f\left(\frac{k(x+1)}{t}\right)dx &= \int_u^\infty (x-u)^{\mu-1}dx\int_0^\infty e^{-k(x+1)y/t}\phi(y)\,dy\\ &= \int_0^\infty e^{-ky/t}\phi(y)\,dy\int_u^\infty e^{-kxy/t}(x-u)^{\mu-1}dx\\ &= \left(\frac{t}{k}\right)^\mu\Gamma(\mu)\int_0^\infty e^{-k(u+1)y/t}y^{-\mu}\phi(y)\,dy\\ &= \left(\frac{t}{k}\right)^\mu\Gamma(\mu)\,f_\mu\left(\frac{k(u+1)}{t}\right),\end{aligned}$$

the interchange of integrations being valid by Fubini's theorem since

$$\begin{aligned}&\int_0^\infty e^{-ky/t}|\phi(y)|\,dy\int_u^\infty e^{-kxy/t}(x-u)^{\mu-1}du\\ &\qquad = \left(\frac{t}{k}\right)^\mu\Gamma(\mu)\int_0^\infty e^{-k(u+1)y/t}y^{-\mu}|\phi(y)|\,dy < \infty.\end{aligned}$$

Thus for $k > \gamma t > 0$,

$$
\begin{aligned}
{}^{\lambda+\mu}L_{k,t}[f(s)] &= \frac{k^{3/2}e^{2k}}{t\pi^{\frac12}}\int_0^\infty x^{\frac12(\lambda+\mu)}J_{\lambda+\mu}(2kx^{\frac12})\,f\!\left(\frac{k(x+1)}{t}\right)dx \\
&= \frac{k^{\mu+3/2}e^{2k}}{t\pi^{\frac12}\Gamma(\mu)}\int_0^\infty f\!\left(\frac{k(x+1)}{t}\right)dx\int_0^x u^{\frac12\lambda}J_\lambda(2ku^{\frac12})(x-u)^{\mu-1}du \\
&= \frac{k^{\mu+3/2}e^{2k}}{t\pi^{\frac12}\Gamma(\mu)}\int_0^\infty u^{\frac12\lambda}J_\lambda(2ku^{\frac12})\,du\int_u^\infty (x-u)^{\mu-1}f\!\left(\frac{k(x+1)}{t}\right)dx \\
&= \frac{t^\mu k^{3/2}e^{2k}}{t\pi^{\frac12}}\int_0^\infty u^{\frac12\lambda}J_\lambda(2ku^{\frac12})\,f_\mu\!\left(\frac{k(u+1)}{t}\right)du \\
&= t^\mu\,{}^{\lambda}L_{k,t}[f_\mu(s)],
\end{aligned}
$$

the interchange of integrations being justified by Fubini's theorem, since by Theorem 2.1 ${}^{\lambda}L_{k,t}[\bar{f}_\mu(s)]$ exists and thus

$$
\begin{aligned}
&\int_0^\infty u^{\frac12\lambda}|J_\lambda(2ku^{\frac12})|\,du\int_u^\infty (x-u)^{\mu-1}\left|f\!\left(\frac{k(x+1)}{t}\right)\right|dx \\
&< \int_0^\infty u^{\frac12\lambda}|J_\lambda(2ku^{\frac12})|\,du\int_u^\infty (x-u)^{\mu-1}\bar{f}\!\left(\frac{k(x+1)}{t}\right)dx \\
&= \left(\frac{t}{k}\right)^\mu\Gamma(\mu)\int_0^\infty u^{\frac12\lambda}|J_\lambda(2ku^{\frac12})|\,\bar{f}_\mu\!\left(\frac{k(u+1)}{t}\right)du < \infty.
\end{aligned}
$$

The theorem is clearly true if $\mu = 0$.

Corollary. *If*

(1) $$e^{-\gamma t}\phi(t) \in L(0, \infty), \quad \gamma > 0.$$

(2) $$t^{-(\frac12\lambda+3/4)}\phi(t) \in L(0, \delta), \textit{ for some } \;> 0,$$

(3) $$\lambda > -1, \; \mu > 0,$$

then $f_{-\mu}(s)$ exists for each $s > \gamma$.

$${}^{\lambda+\mu}L_{k,t}[f_{-\mu}(s)] \text{ and } {}^{\lambda}L_{k,t}[f(s)]$$

exist for each $t > 0$ and all $k > \gamma t$, and for all such k and t,

$${}^{\lambda+\mu}L_{k,t}[f_{-\mu}(s)] = t^\mu\,{}^{\lambda}L_{k,t}[f(s)].$$

Proof. The existence of $f_{-\mu}(s)$ is clear. Let $\psi(t) = t^\mu\phi(t)$ and

$$g(s) = \int_0^\infty e^{-st}\psi(t)\,dt = f_{-\mu}(s).$$

Clearly

$$e^{-\gamma_1 t}\psi(t) \in L(0, \infty)$$

for any $\gamma_1 > \gamma$, and

$$t^{-(\frac12\lambda+\mu+3/4)}\psi(t) = t^{-(\frac12\lambda+3/4)}\phi(t) \in L(0, \delta).$$

Hence ${}^{\lambda+\mu}L_{k,t}[g(s)]$ and $t^\mu\,{}^{\lambda}L_{k,t}[g_\mu(s)]$ exist and are equal, by Theorem 2.2, for each $t > 0$ and all $k > \gamma_1 t$, and hence since γ_1 was an arbitrary number larger than γ, for all $k > \gamma t$. But $g(s) = f_{-\mu}(s)$, $g_\mu(s) = f(s)$, and thus the statement is proved.

Let $I_\alpha\{\phi(u);t\}$ denote the Riemann-Liouville fractional integral of order α, i.e.,

$$I_\alpha\{\phi(u);\, t\} = \begin{cases} (\Gamma(\alpha))^{-1} \displaystyle\int_0^t (t-u)^{\alpha-1}\phi(u)\,du, & \alpha > 0, \\ \phi(t), & \alpha = 0. \end{cases}$$

The next theorem relates the inversion operator and the fractional integral.

THEOREM 2.3. *If*

$$e^{-\gamma t}\phi(t) \in L(0, \infty), \quad \gamma > 0, \tag{1}$$

$$t^{-(\frac{1}{2}\nu+3/4)}\phi(t) \in L(0, \delta), \quad \textit{for some } \delta > 0, \tag{2}$$

$$\nu > -1, \quad \alpha \geq 0, \tag{3}$$

then for each $t > 0$, and all $k > \gamma t$,

$${}^{\nu+\alpha}L_{k,t}[s^{-\alpha}f(s)] \quad \textit{and} \quad I_\alpha\{{}^{\nu}L_{k,u}[f(s)];\, t\}$$

exist and are equal.

Proof. $s^{-\alpha}f(s)$ is the Laplace transform of $I_\alpha\{\phi(u);\, t\}$, which transform exists, by (**4**; ch. 2, §8), for $s > \gamma$. Hence, by Theorem 2.1,

$${}^{\nu+\alpha}L_{k,t}[s^{-\alpha}f(s)]$$

will exist, for each $t > 0$ and all $k > \gamma t$, if $t^{-(\frac{1}{2}(\nu+\alpha)+3/4)}I_\alpha\{\phi(u);t\} \in L(0,\delta)$ for some $\delta > 0$. But, using the same δ as in (2), we have

$$\begin{aligned}
&\int_0^\delta t^{-(\frac{1}{2}(\nu+\alpha)+3/4)}|I_\alpha\{\phi(u);\, t\}|dt \\
&\quad < (\Gamma(\alpha))^{-1}\int_0^\delta t^{-(\frac{1}{2}(\nu+\alpha)+3/4)}dt \int_0^t (t-u)^{\alpha-1}|\phi(u)|\,du \\
&\quad = (\Gamma(\alpha))^{-1}\int_0^\delta t^{-(\frac{1}{2}(\nu+\alpha)+3/4)}dt \int_0^t (t-u)^{\frac{1}{2}\alpha-1}(t-u)^{\frac{1}{2}\alpha}|\phi(u)|\,du \\
&\quad < (\Gamma(\alpha))^{-1}\int_0^\delta t^{-(\frac{1}{2}\nu+3/4)}dt \int_0^t (t-u)^{\frac{1}{2}\alpha-1}|\phi(u)|\,du \\
&\quad = (\Gamma(\alpha))^{-1}\int_0^\delta |\phi(u)|\,du \int_u^\delta (t-u)^{\frac{1}{2}\alpha-1}t^{-(\frac{1}{2}\nu+3/4)}dt \\
&\quad < (\Gamma(\alpha))^{-1}\int_0^\delta u^{-(\frac{1}{2}\nu+3/4)}|\phi\, u)|\,du \int_u^\delta (t-u)^{\frac{1}{2}\alpha-1}dt \\
&\quad = (\tfrac{1}{2}\Gamma(\alpha+1))^{-1}\int_0^\delta u^{-(\frac{1}{2}\nu+3/4)}(\delta-u)^{\frac{1}{2}\alpha}|\phi(u)|\,du \\
&\quad < 2\delta^{\frac{1}{2}\alpha}(\Gamma(\alpha+1))^{-1}\int_0^\delta u^{-(\frac{1}{2}\nu+3/4)}|\phi(u)|\,du < \infty.
\end{aligned}$$

Hence

$${}^{\nu+\alpha}L_{k,t}[s^{-\alpha}f(s)]$$

exists for all $t > 0$ and all $k > \gamma t$.

Now, for $b > \gamma t > 0$,

$$b^{-\alpha}f\left(\frac{b}{t}\right) = t^{1-\alpha}I_\alpha\left\{u^{-(\alpha+1)}f_{-\alpha}\left(\frac{b}{u}\right);\, t\right\}.$$

For

$$
\begin{aligned}
t^{1-\alpha} I_\alpha&\left\{u^{-(\alpha+1)} f_{-\alpha}\left(\frac{b}{u}\right);\, t\right\} \\
&= \frac{t^{1-\alpha}}{\Gamma(\alpha)}\int_0^t (t-u)^{\alpha-1}u^{-(\alpha+1)}du \int_0^\infty e^{-bv/u} v^\alpha \phi(v)\,dv \\
&= \frac{t^{1-\alpha}}{\Gamma(\alpha)}\int_0^\infty v^\alpha\phi(v)\,dv \int_0^t e^{-bv/u}(t-u)^{\alpha-1}u^{-(\alpha+1)}du \\
&= \frac{t^{-\alpha}}{\Gamma(\alpha)}\int_0^\infty e^{-bv/t}v^\alpha\phi(v)\,dv\int_0^\infty e^{-bvx/t}x^{\alpha-1}dx \qquad (\text{where } x=(t/u)-1) \\
&= b^{-\alpha}\int_0^\infty e^{-bv/t}\phi(v)\,dv = b^{-\alpha} f\left(\frac{b}{v}\right).
\end{aligned}
$$

The interchange of integrations is justified by Fubini's theorem, since $b > \gamma t$, and since, with the same change of variables as above,

$$
\frac{t^{1-\alpha}}{\Gamma(\alpha)}\int_0^\infty v^\alpha|\phi(v)|\,dv\int_0^t e^{-bv/u}(t-u)^{\alpha-1}u^{-(\alpha+1)}du = b^{-\alpha}\int_0^\infty e^{-bv/t}|\phi(v)|\,dv < \infty,
$$

by (1).

Thus, setting $b = k(x+1)$, we have

$$
\begin{aligned}
{}^{\nu+\alpha}L_{k,t}[s^{-\alpha}f(s)] &= \frac{k^{3/2}e^{2k}}{t\pi^{\frac12}}\int_0^\infty x^{\frac12(\nu+\alpha)}J_{\nu+\alpha}(2kx^{\frac12})\left(\frac{k(x+1)}{t}\right)^{-\alpha} f\left(\frac{k(x+1)}{t}\right)dx \\
&= \frac{k^{3/2}e^{2k}}{\pi^{\frac12}}\int_0^\infty x^{\frac12(\nu+\alpha)}J_{\nu+\alpha}(2kx^{\frac12})\; I_\alpha\left\{u^{-(\alpha+1)}f_{-\alpha}\left(\frac{k(x+1)}{u}\right);\, t\right\}dx \\
&= I_\alpha\left\{\frac{u^{-\alpha}k^{3/2}e^{2k}}{u\pi^{\frac12}}\int_0^\infty x^{\frac12(\nu+\alpha)}J_{\nu+\alpha}(2kx^{\frac12})\, f_{-\alpha}\left(\frac{k(x+1)}{u}\right)dx;\, t\right\} \\
&= I_\alpha\{u^{-\alpha}\,{}^{\nu+\alpha}L_{k,u}[f_{-\alpha}(s)];\, t\}.
\end{aligned}
$$

The interchange of integrations is justified by Fubini's theorem. For, $k > \gamma t > 0$, and the same argument used to show the existence of ${}^{\nu+\alpha}L_{k,t}[s^{-\alpha}f(s)]$ can be applied to show the existence of ${}^{\nu+\alpha}L_{k,t}[s^{-\alpha}\bar{f}(s)]$, and thus

$$
\begin{aligned}
\frac{k^{3/2}e^{2k}}{\pi^{\frac12}}&\int_0^\infty x^{\frac12(\nu+\alpha)}|J_\nu(2kx^{\frac12})|\; I_\alpha\left\{u^{-(\alpha+1)}\left|f_{-\alpha}\left(\frac{k(x+1)}{u}\right)\right|;\, t\right\}dx \\
&< \frac{k^{3/2}e^{2k}}{\pi^{\frac12}}\int_0^\infty x^{\frac12(\nu+\alpha)}|J_\nu(2kx^{\frac12})|\; I_\alpha\left\{u^{-(\alpha+1)}\bar{f}_{-\alpha}\left(\frac{k(x+1)}{u}\right);\, t\right\}dx \\
&= \frac{k^{3/2}e^{2k}}{\pi^{\frac12}}\int_0^\infty x^{\frac12(\nu+\alpha)}|J_\nu(2kx^{\frac12})|\left(\frac{k(x+1)}{t}\right)^{-\alpha}\bar{f}\left(\frac{k(x+1)}{t}\right)dx < \infty.
\end{aligned}
$$

But by (1), (2), and the corollary to Theorem 2.2,

$$u^{-\alpha}\;{}^{\nu+\alpha}L_{k,u}[f_{-\alpha}(s)] = {}^{\nu}L_{k,u}[f(s)].$$

Hence finally, for each $t > 0$ and all $k > \gamma t$,

$${}^{\nu+\alpha}L_{k,t}[s^{-\alpha}f(s)] = I_\alpha\{{}^{\nu}L_{k,u}[f(s)];\, t\}.$$

The following theorem gives sufficient conditions that ${}^{\nu}L_{k,t}[f(s)]$ exist for a given $f(s)$, and yields a relation involving the Laplace transformation of the operator.

THEOREM 2.4. *If*

(1) $e^{-\gamma u}u^{-1}f(u^{-1}) \in L(0, \infty),\ \gamma > 0,$

(2) $u^{-(\frac{1}{2}\nu+7/4)}f(u^{-1}) \in L(0, \delta),$ *for some* $\delta > 0,$

(3) $g(s) = \displaystyle\int_0^\infty e^{-su}u^{-1}f(u^{-1})\,du,\ s > \gamma,$

(4) $\nu > -1,$

then

(i) ${}^{\nu}L_{k,t}[f(s)]$ *exists for each* $k > 0$ *and almost all* $t > 0,$

(ii) $$\sigma^{-1}\ {}^{\nu}L_{k,\sigma^{-1}}[g(s)] \text{ and } \int_0^\infty e^{-\sigma t}\ {}^{\nu}L_{k,t}[f(s)]dt$$

exist and are equal for each $\sigma > 0$ *and all* $k > \gamma/\sigma.$

Proof. By Theorem 2.1,

$${}^{\nu}L_{k,\sigma^{-1}}[g(s)] \text{ and } {}^{\nu}L_{k,\sigma^{-1}}[\bar{g}(s)]$$

exist for $\sigma > 0$ and all $k > \gamma/\sigma$. But,

$$\begin{aligned}
\sigma^{-1}\ {}^{\nu}L_{k,\sigma^{-1}}[g(s)] &= \frac{k^{3/2}e^{2k}}{\pi^{\frac{1}{2}}}\int_0^\infty x^{\frac{1}{2}\nu}J_\nu(2kx^{\frac{1}{2}})\,dx \int_0^\infty e^{-k(x+1)\sigma u}u^{-1}f(u^{-1})\,du \\
&= \frac{k^{3/2}e^{2k}}{\pi^{\frac{1}{2}}}\int_0^\infty x^{\frac{1}{2}\nu}J_\nu(2kx^{\frac{1}{2}})\,dx \int_0^\infty e^{-\sigma t}t^{-1}f\left(\frac{k(x+1)}{t}\right)dt \text{ (where } t = k(x+1)u) \\
&= \frac{k^{3/2}e^{2k}}{\pi^{\frac{1}{2}}}\int_0^\infty e^{-\sigma t}t^{-1}dt \int_0^\infty x^{\frac{1}{2}\nu}J_\nu(2kx^{\frac{1}{2}})\,f\left(\frac{k(x+1)}{t}\right)dx \\
&= \int_0^\infty e^{-\sigma t}\ {}^{\nu}L_{k,t}\,dt.
\end{aligned}$$

The interchange of integrations is justified by Fubini's theorem, since ${}^{\nu}L_{k,\sigma^{-1}}[\bar{g}(s)]$ exists for $\sigma > 0$ and all $k > \gamma/\sigma$, and hence

$$\begin{aligned}
&\int_0^\infty x^{\frac{1}{2}\nu}|J_\nu(2kx^{\frac{1}{2}})|\,dx \int_0^\infty e^{-k(x+1)\sigma u}u^{-1}|f(u^{-1})|\,du \\
&\qquad = \int_0^\infty x^{\frac{1}{2}\nu}|J_\nu(2kx^{\frac{1}{2}})|\bar{g}(k(x+1)\sigma)\,dx < \infty.
\end{aligned}$$

Thus, by Fubini's theorem ${}^{\nu}L_{k,t}[f(s)]$ exists for each $k > 0$ and almost all $t > 0$, and has a Laplace transform, and the equality stated is true.

COROLLARY. *If*

(1) $e^{-\gamma u}u^{-1}f(u^{-1}) \in L(0, \infty),\ \gamma > 0,$

(2) $g^{\epsilon}(s) = \displaystyle\int_0^\infty e^{-su}u^{-1}e^{-\epsilon/u}f(u^{-1})\,du,\ s > \gamma,\ \epsilon > 0,$

(3) $\nu > -1,$

then

(i) ${}^{\nu}L_{k,t}[e^{-\epsilon s}f(s)]$ *exists for each* $k > 0$ *and almost all* $t > 0$,

(ii) $$\sigma^{-1}\,{}^{\nu}L_{k,\sigma^{-1}}[g^{\epsilon}(s)] \text{ and } \int_0^\infty e^{-\sigma t}\,{}^{\nu}L_{k,t}[e^{-\epsilon s}f(s)]\,dt$$

exist and are equal for each $\sigma > 0$ *and all* $k > \gamma/\sigma$.

Proof. The hypotheses of Theorem 2.4 are clearly fulfilled relative to $e^{-\epsilon s}f(s)$ except possibly (2). But $u^{-(\frac{1}{2}\nu+3/4)}e^{-\epsilon/u}$ is bounded for $u > 0$, and (2) is fulfilled also. Hence, applying the results of Theorem 2.4 to $e^{-\epsilon s}f(s)$, we arrive at the above conclusions.

The next theorem concerns the behaviour of ${}^{\nu}L_{k,t}[e^{-\epsilon s}f(s)]$ when $\phi(t) \in L_p(0, \infty)$, $1 < p < \infty$.

THEOREM 2.5. *If* $\phi(t) \in L_p(0, \infty)$, p *fixed*, $1 < p < \infty$, *and* $\nu > -1$, *then*

(i) ${}^{\nu}L_{k,t}[e^{-\epsilon s}f(s)]$ *exists if* k, t, *and* ϵ *are positive*,

(ii) k_0 *and* N *exist*, N *independent of* k, *such that for* $k > k_0$,

$$\|{}^{\nu}L_{k,t}[e^{-\epsilon s}f(s)]\|_p < N\|\phi(t)\|_p.$$

Proof. Existence follows from the corollary to Theorem 2.1, and for all positive k since any positive γ may be used.

Now ${}^{\nu}L_{k,t}[e^{-\epsilon s}f(s)]$

$$= \frac{k^{3/2}e^{2k}}{t\pi^{\frac{1}{2}}}\int_0^\infty x^{\frac{1}{2}\nu}J_\nu(2kx^{\frac{1}{2}})\,e^{-\epsilon k(x+1)/t}dx\int_0^\infty e^{-k(x+1)u/t}\phi(u)\,du$$

$$= \frac{k^{3/2}e^{2k}}{t\pi^{\frac{1}{2}}}\int_0^\infty e^{-k(u+\epsilon)/t}\phi(u)\,du\int_0^\infty e^{-k(u+\epsilon)x/t}x^{\frac{1}{2}\nu}J_\nu(2kx^{\frac{1}{2}})\,dx$$

$$= \frac{t^{\nu+1}k^{\frac{1}{2}}e^{2k}}{t\pi^{\frac{1}{2}}}\int_0^\infty e^{-k((u+\epsilon)t^{-1}+(u+\epsilon)^{-1}t)}(u+\epsilon)^{-(\nu+1)}\phi(u)\,du,$$

where we have used (**3**; §13.3, (3)), and the interchange of integrations is justified by Fubini's theorem if $k > 0$, $t > 0$, since ${}^{\nu}L_{k,t}[e^{-\epsilon s}\bar{f}(s)]$ exists by the Corollary to Theorem 2.1 applied to $|\phi(t)|$, and thus,

$$\int_0^\infty x^{\frac{1}{2}\nu}|J_\nu(2kx^{\frac{1}{2}})|\,e^{-\epsilon k(x+1)/t}dx\int_0^\infty e^{-k(x+1)u/t}|\phi(u)|\,du < \infty.$$

Hence, for $1 < p < \infty$, we have, using Hölder's inequality,

$$|{}^{\nu}L_{k,t}[e^{-\epsilon s}f(s)]|^p$$

$$< \left(\frac{t^{\nu}k^{\frac{1}{2}}e^{2k}}{\pi^{\frac{1}{2}}}\right)^p\int_0^\infty e^{-k((u+\epsilon)t^{-1}+(u+\epsilon)^{-1}t)}(u+\epsilon)^{-(\nu+1)}|\phi(u)|^p du$$

$$\left\{\int_0^\infty e^{-k(ut^{-1}-tu^{-1})}u^{-(\nu+1)}du\right\}^{p/q}$$

$$< \left\{\frac{2k^{\frac{1}{2}}e^{k}\,K_\nu(2k)}{\pi^{\frac{1}{2}}}\right\}^{p/q}\frac{t^{\nu}k^{\frac{1}{2}}e^{2k}}{\pi^{\frac{1}{2}}}\int_0^\infty e^{-k((u+\epsilon)t^{-1}+(u+\epsilon)^{-1}t)}(u+\epsilon)^{-(\nu+1)}|\phi(u)|^p du.$$

by (**3**; §6.23, (8)), where $q = p/(p-1)$. Hence

$$\int_0^\infty |{}^{\nu}L_{k,t}[e^{-\epsilon s}f(s)]|^p dt$$
$$< \left\{\frac{2k^{\frac{1}{2}}e^{2k}K_\nu(2k)}{\pi^{\frac{1}{2}}}\right\}^{p/q} \frac{k^{\frac{1}{2}}e^{2k}}{\pi^{\frac{1}{2}}} \int_0^\infty |\phi(u)|^p du \int_0^\infty e^{-k((u+\epsilon)t^{-1}+(u+\epsilon)^{-1}t)} t^\nu (u+\epsilon)^{-(\nu+1)} dt$$
$$= \left\{\frac{2k^{\frac{1}{2}}e^{2k}K_\nu(2k)}{\pi^{\frac{1}{2}}}\right\}^{p/q} \frac{2k^{\frac{1}{2}}e^{2k}K_{\nu+1}(2k)}{\pi^{\frac{1}{2}}} \int_0^\infty |\phi(u)|^p du.$$

Hence

$$||{}^{\nu}L_{k,t}[e^{-\epsilon s}f(s)]||_p < \left\{\frac{2k^{\frac{1}{2}}e^{2k}K_\nu(2k)}{\pi^{\frac{1}{2}}}\right\}^{1/q} \left\{\frac{2k^{\frac{1}{2}}e^{2k}K_{\nu+1}(2k)}{\pi^{\frac{1}{2}}}\right\}^{1/p} ||\phi(t)||_p.$$

But by (3; §7.23, (1)), $K_\lambda(2k) \sim e^{-2k}\pi^{\frac{1}{2}}/2k^{\frac{1}{2}}$ as $k \to \infty$. Hence $k_0(\lambda)$ and N exist such that for $k > k_0$, $2k^{\frac{1}{2}}e^{2k}K_\lambda(2k)\pi^{-\frac{1}{2}} < N$. Thus if $k_0 \geqslant \max(k_0(\nu), k_0(\nu+1))$, we have

$$||{}^{\nu}L_{k,t}[e^{-\epsilon s}f_{(s)}]||_p < N||\phi(t)||_p, \qquad k > k_0.$$

If $p = \infty$, we have

$$|{}^{\nu}L_{k,t}[e^{-\epsilon s}f(s)]| < \frac{k^{\frac{1}{2}}e^{2k}}{\pi^{\frac{1}{2}}} \int_0^\infty e^{-k(ut^{-1}+tu^{-1})} t^\nu u^{-(\nu+1)} du ||\phi(t)||_\infty$$
$$= \frac{2k^{\frac{1}{2}}e^{2k}K_\nu(2k)}{\pi^{\frac{1}{2}}} ||\phi(t)||_\infty < N||\phi(t)||_\infty,$$

for $k > k_0(\nu)$, so that

$$||{}^{\nu}L_{k,t}[e^{-\epsilon s}f(s)]||_\infty < N||\phi(t)||_\infty.$$

Corollary. *If $\phi(t) \in L_p(0, \infty)$, p fixed, $1 < p < \infty$, $p > 4/(1-2\nu)$, and $\nu > -1$, then*

(1) *${}^{\nu}L_{k,t}[f(s)]$ exists if k and t are positive,*

(2) *k_0 and N exist, N independent of k, such that for $k > k_0$,*

$$||{}^{\nu}L_{k,t}[f(s)]||_p < N||\phi(t)||_p.$$

Proof. If $p > 4/(1-2\nu)$, then $q(\frac{1}{2}\nu + 3/4) < 1$, where $q = p/(p-1)$. Hence, by Hölder's inequality,

$$\int_0^\delta t^{-(\frac{1}{2}\nu+3/4)}|\phi(t)|\, dt < \left\{\int_0^\delta t^{-q(\frac{1}{2}\nu+3/4)} dt\right\}^{1/q} \left\{\int_0^\delta |\phi(t)|^p dt\right\}^{1/p} < \infty,$$

so that, by Theorem 2.1, ${}^{\nu}L_{k,t}[f(s)]$ exists for $t > 0$, and for all $k > 0$ since any $\gamma > 0$ may be used.

The remainder of the corollary follows in exactly the same manner as in Theorem 2.5.

3. Inversion of the transformation. The following three theorems comprise the inversion theory for the operator.

Theorem 3.1. *If*

(1) $e^{-\gamma t}\phi(t) \in L(0, \infty)$, $\gamma > 0$,

and if

(2) $\nu \geqslant -\frac{1}{2}$ *and* $t^{-(\nu+1)}\phi(t) \in L(0, \delta)$ *for some* $\delta > 0$,

or if

(2′) $-1 < \nu < \frac{1}{2}$ *and* $t^{-(\frac{1}{2}\nu+3/4)}\phi(t) \in L(0, \delta)$ *for some* $\delta > 0$, *then, at each point* $t > 0$ *of the Lebesgue set of* ϕ,

$$\lim_{k\to\infty} {}^{\nu}L_{k,t}[f(s)] = \phi(t).$$

Proof. Case (a): $\nu \geqq -\frac{1}{2}$. Setting $\lambda = -\frac{1}{2}$ and $\mu = \nu + \frac{1}{2}$ in Theorem 2.2, we have that ${}^{\nu}L_{k,t}[f(s)]$ and $t^{(\nu+\frac{1}{2})}\, {}^{-\frac{1}{2}}L_{k,t}[f_{\nu+\frac{1}{2}}(s)]$ exist and are equal. But by (**2**; Theorem 1), under the above conditions

$$\lim_{k\to\infty} {}^{-\frac{1}{2}}L_{k,t}[f_{\nu+\frac{1}{2}}(s)] = t^{-(\nu+\frac{1}{2})}\, \phi(t)$$

at each point $t > 0$ of the Lebesgue set of ϕ. Hence, at each such point,

$$\lim_{k\to\infty} {}^{\nu}L_{k,t}[f(s)] = \phi(t).$$

Case (b): $-1 < \nu < \frac{1}{2}$. Setting $\lambda = \nu$, $\mu = -(\nu + \frac{1}{2})$ in the Corollary to Theorem 2.2, we have that ${}^{-\frac{1}{2}}L_{k,t}[f_{\nu+\frac{1}{2}}(s)]$ and $t^{-(\nu+\frac{1}{2})}\, {}^{\nu}L_{k,t}[f(s)]$ are equal. Hence, as in case (a), using (**2**; Theorem 1) we have

$$\lim_{k\to\infty} {}^{\nu}L_{k,t}[f(s)] = \phi(t)$$

at each point $t > 0$ of the Lebesgue set of ϕ.

THEOREM 3.2. *If*
(1) $e^{-\gamma t}\phi(t) \in L(0, \infty)$, $\gamma > 0$,
(2) $\alpha > 0$, $\nu > -1$,
and if
(3) $\alpha + \nu \geqq -\frac{1}{2}$ *and* $t^{-(\nu+1)}\phi(t) \in L(0, \delta)$ *for some* $\delta > 0$,
or if
(3′) $\alpha + \nu < -\frac{1}{2}$ *and* $t^{-(\frac{1}{2}\nu+3/4)}\phi(t) \in L(0, \delta)$ *for some* $\delta > 0$,
then for each $t > 0$, *of the Lebesgue set of* $I_\alpha\{\phi(u); t\}$

$$\lim_{k\to\infty} I_\alpha\{{}^{\nu}L_{k,u}[f(s)]; t\} = I_\alpha\{\phi(u); t\}.$$

Proof. By Theorem 2.3, under the above conditions

$$I_\alpha\{{}^{\nu}L_{k,u}[f(s)]; t\} = {}^{\nu+\alpha}L_{k,t}[s^{-\alpha}f(s)].$$

Thus the theorem will follow from Theorem 3.1, if either $\nu + \alpha \geqq -\frac{1}{2}$ and $t^{-(\nu+\alpha+1)}I_\alpha\{\phi(u); t\} \in L(0, \delta)$ for some $\delta > 0$, or $\nu + \alpha < -\frac{1}{2}$ and $t^{-(\frac{1}{2}(\nu+\alpha)+3/4)}$ $I_\alpha\{\phi(u); t\} \in L(0, \delta)$ for some $\delta > 0$. But if $\nu + \alpha \geqq -\frac{1}{2}$, then $t^{-(\nu+1)}$ $\phi(t) \in L(0, \delta)$, and thus

$$\begin{aligned}
\int_0^\delta t^{-(\nu+\alpha+1)}|I_\alpha\{\phi(u); t\}|\, dt &\leqq (\Gamma(\alpha))^{-1}\int_0^\delta t^{-(\nu+\alpha+1)}dt \int_0^t (t-u)^{\alpha-1}|\phi(u)|\, du \\
&= (\Gamma(\alpha))^{-1}\int_0^\delta |\phi(u)|\, du \int_u^\delta t^{-(\nu+\alpha+1)}(t-u)^{\alpha-1}dt \\
&= (\Gamma(\alpha))^{-1}\int_0^\delta u^{-(\nu+1)}|\phi(u)|\, du \int_{u/\delta}^1 v^{\nu}(1-v)^{\alpha-1}dv \qquad (\text{where } v = u/t) \\
&\leqq (\Gamma(\alpha))^{-1}\int_0^\delta u^{-(\nu+1)}|\phi(u)|\, du \int_0^1 v^{\nu}(1-v)^{\alpha-1}dv < \infty.
\end{aligned}$$

Further, we have already shown in the course of the proof of Theorem 2.3 that if $t^{-(\frac{1}{2}\nu+3/4)}\phi(t) \in L(0, \delta)$, then so is $t^{-(\frac{1}{2}(\nu+\alpha)+3/4)}I_\alpha\{\phi(u); t\}$.

We conclude this section with a theorem removing the restrictions on the behaviour of $\phi(t)$ at $t = 0$.

THEOREM 3.3. *If*
(1) $e^{-\gamma t}\phi(t) \in L(0, \infty)$, $\gamma > 0$,
(2) $\nu > -1$.
then for each $\epsilon > 0$, *and at each point* $t > \epsilon$ *such that* $t - \epsilon$ *is in the Lebesgue set of* $\phi(u)$, *we have*

$$\lim_{k\to\infty} {}^\nu L_{k,t}[e^{-\epsilon s}f(s)] = \phi(t-\epsilon).$$

Proof. We have $e^{-\epsilon s}f(s) = \int_0^\infty e^{-st}\phi_\epsilon(t)\,dt$, where $\phi_\epsilon(t) = 0$, $0 < t < \epsilon$, and $\phi_\epsilon(t) = \phi(t-\epsilon)$, $t > \epsilon$. Thus the hypotheses of Theorem 3.1 are clearly fulfilled relative to $\phi_\epsilon(t)$, and the conclusion follows.

4. Representation theorems. The first theorem of this section is fundamental in the representation theory.

THEOREM 4.1. *If*
(1) $e^{-\gamma u}u^{-1}f(u^{-1}) \in L(0, \infty)$, $\gamma > 0$,
and if
(2) $\nu > -\frac{1}{2}$, *and* $u^{-(\nu+2)}f(u^{-1}) \in L(0, \delta)$ *for some* $\delta > 0$, *or if*
(2′) $-1 < \nu < -\frac{1}{2}$, *and* $u^{-(\frac{1}{2}\nu+7/4)}f(u^{-1}) \in L(0, \delta)$ *for some* $\delta > 0$,
then ${}^\nu L_{k,t}[f(s)]$ *exists for each* $k > 0$ *and almost all* $t > 0$, ${}^\nu L_{k,t}[f(s)]$ *has a Laplace transform, and*

$$\lim_{k\to\infty} \int_0^\infty e^{-\sigma t}\, {}^\nu L_{k,t}[f(s)]\,dt = f(\sigma),$$

at each point $\sigma > 0$ *of the Lebesgue set of* f.

Proof. The existence of ${}^\nu L_{k,t}[f(s)]$ follows from Theorem 2.4, as does the existence of its Laplace transform for each $\sigma > 0$ and all $k > \gamma/\sigma$. The remainder of the theorem follows from Theorems 2.4 and 3.1.

COROLLARY. *If*
(1) $e^{-\gamma u}u^{-1}f(u^{-1}) \in L(0, \infty)$, $\gamma > 0$,
(2) $\nu > -1$, $\epsilon > 0$,
then ${}^\nu L_{k,t}[e^{-\epsilon s}f(s)]$ *exists for each* $k > 0$ *and almost all* $t > 0$, ${}^\nu L_{k,t}[e^{-\epsilon s}f(s)]$ *has a Laplace transform, and*

$$\lim_{k\to\infty} \int_0^\infty e^{-\sigma t}\, {}^\nu L_{k,t}[e^{-\epsilon s}f(s)]\,dt = e^{-\epsilon\sigma}f(\sigma),$$

at each point $\sigma > 0$ *of the Lebesgue set of* f.

Proof. Since $u^{-(\nu+1)}e^{-\epsilon/u}$ and $u^{-(\frac{1}{2}\nu+7/4)}e^{-\epsilon/u}$ are uniformly bounded in u for $u > 0$, the hypotheses of Theorem 4.1 are satisfied, and the conclusion follows.

THEOREM 4.2. *Necessary and sufficient conditions that $f(s)$ be equal almost everywhere to the Laplace transform of a function of the form $t^{\alpha}\phi(t)$, where $\phi(t) \in L_p(0, \infty)$, p fixed, $1 < p \leq \infty$ and $\alpha > -(p-1)/p$, are that*
(1) $e^{-\gamma u}u^{-1}f(u^{-1}) \in L(0, \infty), \gamma > 0$,
(2) $\|t^{-\alpha}\, {}^{\nu+\alpha}L_{k,t}[e^{-\epsilon s}f(s)]\|_p \leq M_p$,
where M_p is independent of ϵ and k, for all $\epsilon > 0$ and $k > k_0$, for some $\nu > \max(-1, -1-\alpha)$.

Proof of necessity. Let

$$f(s) = \int_0^{\infty} e^{-st}t^{\alpha}\phi(t)\, dt,$$

where $\phi(t) \in L_p(0, \infty)$, p fixed, $1 < p \leq \infty$, and $\alpha > -(p-1)/p$.

If $p < \infty$, and $q = p/(p-1)$, we have by Hölder's inequality, for $s > 0$,

$$|f(s)| \leq \left\{ \int_0^{\infty} |\phi(t)|^p dt \right\}^{1/p} \left\{ \int_0^{\infty} e^{-qst}t^{\alpha q} dt \right\}^{1/q}$$
$$= (qs)^{-(\alpha+1/q)}(\Gamma(1+\alpha q))^{1/q}\|\phi(t)\|_p = Ms^{-(\alpha+1/q)}.$$

Thus, for any $\gamma > 0$,

$$e^{-\gamma u}u^{-1}|f(u^{-1})| \leq Me^{-\gamma u}u^{-(1-\alpha-1/q)} = Me^{-\gamma u}u^{\alpha-1/p},$$

and hence since $\alpha - 1/p > -p/p = -1$, we find that $e^{-\gamma u}u^{-1}f(u^{-1}) \in L(0, \infty)$.

If $p = \infty$, we have for $s > 0$,

$$|f(s)| \leq \int_0^{\infty} e^{-st}t^{\alpha}dt\|\phi(t)\|_{\infty} = s^{-(\alpha+1)}\Gamma(\alpha+1)\|\phi(t)\|_{\infty},$$

and hence for any $\gamma > 0$, since $\alpha > -1$,

$$e^{-\gamma u}u^{-1}|f(u^{-1})| \leq e^{-\gamma u}u^{\alpha}\Gamma(\alpha+1)\|\phi(t)\|_{\infty} \in L(0, \infty).$$

For the necessity of (2), we have by Theorem 2.2, if $\alpha \geq 0$,

$$t^{-\alpha}\, {}^{\nu+\alpha}L_{k,t}[e^{-\epsilon s}f(s)] = {}^{\nu}L_{k,t}[(e^{-\epsilon s}f(s))_{\alpha}] = {}^{\nu}L_{k,t}[e^{-\epsilon s}g(s)]$$

where

$$g(s) = \int_0^{\infty} e^{-st}t^{\alpha}(t+\epsilon)^{-\alpha}\phi(t)\, dt$$

If $\alpha < 0$, we set $\lambda = \nu + \alpha$, $\mu = -\alpha$ in the corollary to Theorem 2.2 and obtain again

$$t^{-\alpha}\, {}^{\nu+\alpha}L_{k,t}[e^{-\epsilon s}f(s)] = {}^{\nu}L_{k,t}[e^{-\epsilon s}g(s)].$$

But

$$g(s) = \int_0^{\infty} e^{-st}t^{\alpha}(t+\epsilon)^{-\alpha}\phi(t)dt$$

is the Laplace transform of a function in L_p. In fact, since $t^{\alpha}(t+\epsilon)^{-\alpha} < 1$ for $t \geq 0$, we have $\|t^{\alpha}(t+\epsilon)^{-\alpha}\phi(t)\|_p \leq \|\phi(t)\|_p < \infty$. Hence we may apply Theorem 2.5, and obtain, for $k > k_0$,

$$\|t^{-\alpha}\, {}^{\nu+\alpha}L_{k,t}[e^{-\epsilon s}f(s)]\|_p = \|{}^{\nu}L_{k,t}[e^{-\epsilon s}g(s)]\|_p \leq N\|t^{\alpha}(t+\epsilon)^{-\alpha}\phi(t)\|_p \leq N\|\phi(t)\|_p.$$

Proof of sufficiency. By (2) and (**4**; ch. 1, Theorem 17a for $1 < p < \infty$, Theorem 17b for $p = \infty$), there exists, for each $\epsilon > 0$, an increasing unbounded sequence $\{k_i(\epsilon)\}$, and a function $\phi_\epsilon(t) \in L_p(0, \infty)$, with $||\phi_\epsilon(t)|| < M_p$, such that

$$\lim_{i \to \infty} \int_0^\infty \alpha(t)\, t^{-\alpha}\, {}^{\nu+\alpha}L_{k_i, t}[e^{-\epsilon s} f(s)]\, dt = \int_0^\infty \alpha(t)\, \phi_\epsilon(t)\, dt$$

for every function $\alpha(t) \in L_q(0, \infty)$. But for each $\sigma > 0$, $e^{-\sigma t} t^\alpha \in L_q(0, \infty)$, since $\alpha > -1/q$. Hence, by the Corollary to Theorem 4.1, we have for almost all $\sigma > 0$

$$\begin{aligned} e^{-\epsilon\sigma} f(\sigma) &= \lim_{i \to \infty} \int_0^\infty e^{-\sigma t}\, {}^{\nu+\alpha}L_{k_i, t}[e^{-\epsilon s} f(s)]\, dt \\ &= \lim_{i \to \infty} \int_0^\infty e^{-\sigma t} t^\alpha\, t^{-\alpha}\, {}^{\nu+\alpha}L_{k_i, t}[e^{-\epsilon s} f(s)]\, dt = \int_0^\infty e^{-\sigma t} t^\alpha \phi_\epsilon(t)\, dt. \end{aligned}$$

But by the same theorems of (**4**), and since $||\phi_\epsilon(t)||_p < M_p$, there exists a sequence $\{\epsilon_i\}$, with $\lim \epsilon_i = 0$, and a function $\phi(t) \in L_p(0, \infty)$ such that $||\phi(t)|| < M_p$, and such that for every $\alpha(t) \in L_q(0, \infty)$,

$$\lim_{i \to \infty} \int_0^\infty \alpha(t)\, \phi_{\epsilon_i}(t)\, dt = \int_0^\infty \alpha(t)\, \phi(t)\, dt.$$

Let Σ_ϵ be the set of measure zero in which, for $\sigma > 0$,

$$e^{-\epsilon\sigma} f(\sigma) \neq \int_0^\infty e^{-\sigma t} t^\alpha \phi_\epsilon(t)\, dt;$$

let

$$\Sigma = \bigcup_{i=1}^\infty \Sigma_{\epsilon_i}.$$

Then Σ has measure zero.

Let σ be positive and $\sigma \notin \Sigma$. Then since $e^{-\sigma t} t^\alpha \in L_q(0, \infty)$, we have

$$f(\sigma) = \lim_{i \to \infty} e^{-\epsilon_i \sigma} f(\sigma) = \lim_{i \to \infty} \int_0^\infty e^{-\sigma t} t^\alpha \phi_{\epsilon_i}(t)\, dt = \int_0^\infty e^{-\sigma t} t^\alpha \phi(t)\, dt.$$

and the theorem is proved.

Corollary. *Necessary and sufficient conditions that $f(s)$ be equal almost everywhere to the Laplace transform of a function in $L_p(0, \infty)$, p fixed, $1 < p < \infty$, are that*

(1) $e^{-\gamma u} u^{-1} f(u^{-1}) \in L(0, \infty)$, $\gamma > 0$,

(2) $||{}^\nu L_{k,t}[e^{-\epsilon s} f(s)]||_p < M_p$, *for all* $\epsilon > 0$ *and* $k > k_0$, *and for some* $\nu > -1$.

Proof. Since $(p-1)/p > 0$, we may always set $\alpha = 0$ in Theorem 4.2 and the result follows.

The next theorem shows how far the factor $e^{-\epsilon s}$ is necessary.

Theorem 4.3. *Sufficient conditions that $f(s)$ be equal to the Laplace transformation of a function of the form $t^\alpha \phi(t)$, $\phi(t) \in L_p(0, \infty)$, p fixed, $1 < p < \infty$, $\alpha > -(p-1)/p$, are that*

(1) $e^{-\gamma u}u^{-1}f(u^{-1}) \in L(0, \infty)$, $\gamma > 0$,
and
(2) $u^{-(\nu+\alpha+2)}f(u^{-1}) \in L(0, \delta)$ *for some* $\delta > 0$, *and some* $\nu > -1$ *such that* $\nu + \alpha \geqslant -\frac{1}{2}$,
or
(2′) $u^{-(\frac{1}{2}(\nu+\alpha)+7/4)}f(u^{-1}) \in L(0, \delta)$ *for some* $\delta > 0$, *and some* ν *such that* $\nu > -1$, *and* $-1 < \nu + \alpha < -\frac{1}{2}$,
(3) $\|t^{-\alpha}\,{}^{\nu+\alpha}L_{k,t}[f(s)]\|_p < M_p$, $k > k_0$.

Condition (1) *is necessary for all such* p *and* α. *Condition* (2) *is necessary if* $p\nu > -1$, *and* (2′) *is necessary if* $p(1 + 2(\alpha - \nu)) > 4$. *Condition* (3) *is necessary if* $p > 4/(1 - 2\nu)$.

Proof of necessity. Let
$$f(s) = \int_0^\infty e^{-st}t^\alpha \phi(t)\, dt$$
where $\phi(t) \in L_p(0, \infty)$, p fixed, $1 < p < \infty$, and $\alpha > -(p-1)/p$.

The proof of the necessity of (1) is the same as in the preceding theorem, and, as in its proof, we have $|f(s)| < Ms^{-(\alpha+1/q)}$ where $q = p/(p-1)$. Hence
$$u^{-(\nu+\alpha+2)}|f(u^{-1})| < Mu^{-(\nu+1+1/p)},$$
and thus (2) is necessary if $\nu + 1/p > 0$, i.e., if $p\nu > -1$. Also,
$$u^{-(\frac{1}{2}(\nu+\alpha)+7/4)}|f(u^{-1})| < Mu^{-(\frac{1}{2}(\nu-\alpha)+1/p+\frac{3}{4})}$$
so that (2′) is necessary if $\frac{1}{2}(\nu - \alpha) + 1/p + \frac{3}{4} < 1$, i.e., if $p(1 + 2(\alpha - \nu)) > 4$.

For the necessity of (3), we have, since $p > 4/(1 - 2\nu)$ that $q(\frac{1}{2}\nu + \frac{3}{4}) < 1$. Hence if $\alpha \geqslant 0$, the hypotheses of Theorem 2.2 are fulfilled, and $t^{-\alpha}\,{}^{\nu+\alpha}L_{k,t}[f(s)] = {}^{\nu}L_{k,t}[f_\alpha(s)]$. Thus by the Corollary to Theorem 2.4, since $f_\alpha(s)$ is the Laplace transform of $\phi(t)$, and $\phi(t) \in L_p(0, \infty)$,
$$\|t^{-\alpha}\,{}^{\nu+\alpha}L_{k,t}[f(s)]\|_p < N\|\phi(t)\|_p.$$

If $\alpha < 0$, and if we set $\lambda = \nu + \alpha$, $\mu = -\alpha$, the hypotheses of the Corollary to Theorem 2.2 are fulfilled and the same results ensue. Hence (3) is necessary.

Proof of sufficiency. By (3) and (**4**; ch. 1, Theorem 17a for $1 < p < \infty$, Theorem 17b for $p = \infty$) there exists an increasing unbounded sequence $\{k_i\}$ and a function $\phi(t) \in L_p(0, \infty)$, with $\|\phi(t)\|_p < M_p$, and such that
$$\lim_{i\to\infty} \int_0^\infty \alpha(t)\, t^{-\alpha}\,{}^{\nu+\alpha}L_{k_i,t}[f(s)]\, dt = \int_0^\infty \alpha(t)\, \phi(t)\, dt$$
for every function $\alpha(t) \in L_q(0, \infty)$. But for each $\sigma > 0$, $e^{-\sigma t}t^\alpha \in L_q(0, \infty)$ since $\alpha > -1/q$. Thus, by Theorem 4.1, we have for almost all $\sigma > 0$,
$$\begin{aligned} f(\sigma) &= \lim_{i\to\infty} \int_0^\infty e^{-\sigma t}\,{}^{\nu}L_{k_i,t}[f(s)]\, dt \\ &= \lim_{i\to\infty} \int_0^\infty e^{-\sigma t}t^\alpha(t^{-\alpha}\,{}^{\nu}L_{k_i,t}[f(s)])\, dt \\ &= \int_0^\infty e^{-\sigma t}t^\alpha \phi(t)\, dt. \end{aligned}$$

References

1. A. Erdélyi, *The inversion of the Laplace integral*, Math. Mag., *24* (1950–51), 1–7.
2. P. G. Rooney, *A new inversion and representation theory for the Laplace transformation*, Can. J. Math., *4* (1953), 436–444.
3. G. N. Watson, *A treatise on Bessel functions* (2nd. ed., New York, 1941).
4. D. V. Widder, *The Laplace Transform* (Princeton, 1941).

University of Alberta, and
Summer Research Institute

ASYMPTOTIC BEHAVIOUR OF THE INVERSE OF A LAPLACE TRANSFORM

T. E. HULL AND C. FROESE

1. Introduction. There are many problems, particularly in circuit theory, where the inverse of a Laplace transform is required but only for large values of the independent variable. For example (**2**, p. 66) the Laplace transform

$$i(s) = \int_0^\infty e^{-st} I(t)\, dt$$

of the current function for a semi-infinite cable under certain conditions turns out to be

$$i(s) = \frac{s^{\frac{1}{2}}}{1+s}.$$

To find the behaviour of $I(t)$ for large t we can proceed as follows. We first expand $i(s)$ in powers of s to obtain

$$i(s) = s^{\frac{1}{2}} - s^{3/2} + s^{5/2} - \ldots .$$

We then find the correct asymptotic expansion

$$I(t) \sim -\left(\frac{2}{\pi}\right)^{\frac{1}{2}} [(2t)^{-3/2} + 1{\cdot}3(2t)^{-5/2} + 1{\cdot}3{\cdot}5(2t)^{-7/2} + \ldots]$$

by using *formally*, for all α, the result

$$\int_0^\infty e^{-st} \frac{t^{\alpha-1}}{\Gamma(\alpha)}\, dt = s^{-\alpha}$$

which actually holds only for $\Re(\alpha) > 0$.

Heaviside (**6**) developed this procedure in a somewhat different form and used it to solve a large number of practical problems. But it was only many years later with the development of Laplace transform theory that his methods were established on a rigorous basis (**1**; **2**).

More recently Doetsch (**4**) has given a number of rather loosely connected theorems which enable one to handle more complicated transforms than the one given above. He discusses in detail the cases when the transform function behaves near the origin like $s^{-\alpha}\log^n s$, all α and $n = 0, 1, 2, \ldots$.

It is the purpose of this paper to show how his methods can be improved and generalized so as to give results when the transform function behaves near the origin like, for example, $s^{-\alpha}e^{-\beta/s}$ or $s^{-\alpha}/\log s$, all α, β.

The main results are contained in two theorems—one about properties of the inversion integral and one about comparison functions. Some of the for-

Received May 31, 1954. This investigation was carried out while Miss Froese held a National Research Council Scholarship.

mulas needed for applications are collected in a table and a few examples are treated briefly at the end.

2. Properties of the inversion integral. The Laplace transform $f(s)$ of a function $F(t)$ is defined in the right half plane $\Re(s) > s_0$ by

$$f(s) = \int_0^\infty e^{-st}F(t)\,dt$$

if this integral exists for $\Re(s) > s_0$. Under some circumstances the inversion integral

$$F(t) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} e^{st}f(s)\,ds, \qquad t > 0,$$

is valid when $a > s_0$. In what follows all such integrals will be considered Lebesgue integrals or principal values of Lebesgue integrals. For example, the inversion integral is to be taken as a principal value so that

$$\int_{a-i\infty}^{a+i\infty} \text{ stands for } \lim_{R\to\infty}\int_{a-iR}^{a+iR}.$$

We shall begin with a theorem about certain properties of the inversion integral. Our plan is to first impose a few mild restrictions on the transform function. To avoid unnecessary complications we restrict ourselves to considering singularities only at the origin. We then show that the path of integration can be altered in such a way that the asymptotic behaviour of the integral is determined by its contribution from a small portion of the contour around the origin. Obvious generalizations will appear when we consider examples in §4.

THEOREM 1. *If*

(i) *$f(s)$ is an analytic function of s, regular for $|\arg s| < \psi$ for some ψ where $\frac{1}{2}\pi < \psi < \pi$, except perhaps for a singularity at the origin,*[1]

(ii) *$e^{ks}f(s) \to 0$ uniformly as $|s| \to \infty$ in the sectors $\frac{1}{2}\pi < |\arg s| < \psi$, for some finite k,*

(iii) $$F(t) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} e^{st}f(s)\,ds$$

exists for $t > 0$, at least as a principal value, for some finite $a > 0$, and

(iv) *$f(s) \to 0$ as $|s| \to \infty$ in the strip $0 < \Re(s) < a$, then*

$$F(t) = \frac{1}{2\pi i}\int_C e^{st}f(s)\,ds, \qquad t > k, \tag{1}$$

where C is the heavy contour shown in Fig. 1 *and where the integral is taken to be a principal value; moreover,*

$$\frac{1}{2\pi i}\int_{C'} e^{st}f(s)\,ds = o(e^{-\epsilon t}), \qquad t \to \infty, \tag{2}$$

for some $\epsilon > 0$, where C' is the part of C for which $|s| > r > 0$.

[1]In case $f(s)$ is not single-valued we shall restrict our attention to one of its branches.

For the proof we notice first that, because of (i), the integral around the *closed* contour in Fig. 1 is zero. The first result then follows from (iii) as soon as we can show that the contribution when $|s| = R$ to the integral around the closed contour approaches zero as $R \to \infty$ while $t > k$.

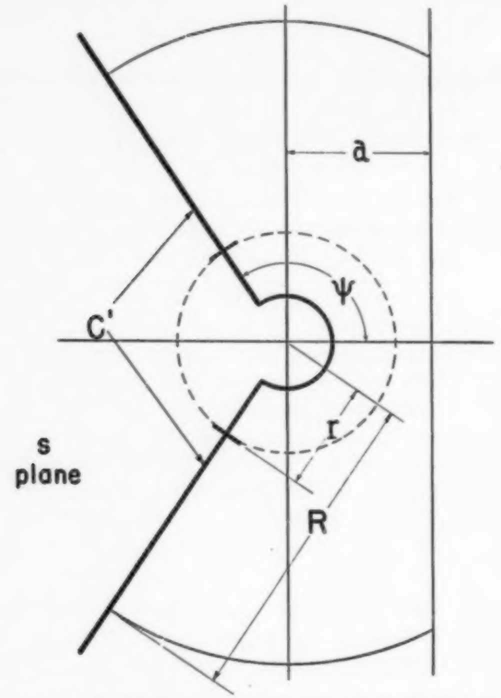

Fig. 1. C is the heavy contour. C' is the portion of C with $|s| > r > 0$.

The parts of this contribution in the strip $0 < \Re(s) < a$ approach zero because of (iv); the remaining parts approach zero because, using (ii), the absolute value I of the part in the second quadrant is

$$I < \frac{R\delta}{2\pi} \int_{\pi/2}^{\psi} e^{Rt\cos\theta} e^{-kR\cos\theta}\, d\theta$$

where $\delta \to 0$ as $R \to \infty$. Hence

$$\begin{aligned} I &< \frac{\delta R}{2\pi} \int_{\pi/2}^{\psi} e^{R(t-k)(1-2\theta/\pi)}\, d\theta \\ &= \frac{\delta}{4(t-k)} \left[1 - e^{R(t-k)(1-2\psi/\pi)}\right] \\ &\to 0 \qquad\qquad R \to \infty \text{ and } t > k, \end{aligned}$$

and similarly for the part in the third quadrant. The second result follows because

$$\left|\frac{1}{2\pi i}\int_{C'} e^{st} f(s)\,ds\right| < \frac{\delta}{\pi}\left|\int_r^\infty e^{(t-k)u\cos\psi}\,du\right|$$

$$= \frac{\delta e^{-kr\cos\psi}}{-(t-k)\cos\psi}\,(e^{tr\cos\psi})$$

$$= o(e^{-\epsilon t}) \qquad \text{as } t \to \infty,$$

for $\epsilon = -r\cos\psi > 0$.

3. Comparison functions. We shall now consider a theorem about certain standard transforms which will be used later for comparison with a given transform. We shall prove

THEOREM 2. *If*

$$F(t) = \begin{cases} 0\ , & 0 < |t| < 1, \\ \frac{1}{2}\Phi(t), & |t| = 1, \\ \Phi(t), & |t| > 1, \end{cases}$$

where $\Phi(t)$ is any function from the second column of the table, then the Laplace transform of $F(t)$ exists for $\Re(s) > 0$ and can be represented for all s by

$$f(s) = \phi(s) + \textit{an entire function},$$

where $\phi(s)$ is the function from the table corresponding to the $\Phi(t)$ being considered. Moreover $f(s)$ satisfies the conditions of Theorem 1 (*and the integral defined in* (iii) *is the $F(t)$ given above*).

We shall first prove the theorem for case 1 of the table and then show how this result can be generalized to give the remaining results.

Case 1. Here

$$\Phi(t) = \frac{t^{\alpha-1}}{\Gamma(\alpha)}$$

so that $f(s)$ exists and

$$f(s) = \int_1^\infty e^{-st}\frac{t^{\alpha-1}}{\Gamma(\alpha)}\,dt, \qquad \Re(s) > 0. \tag{3}$$

For $\Re(\alpha) > 0$ we can write

$$f(s) = \int_0^\infty e^{-st}\frac{t^{\alpha-1}}{\Gamma(\alpha)}\,dt - \int_0^1 e^{-st}\frac{t^{\alpha-1}}{\Gamma(\alpha)}\,dt \tag{4}$$

$$= s^{-\alpha} + \text{an entire function}, \qquad \Re(s) > 0.$$

When $\Re(\alpha) < 0$ and $\alpha \neq 0, -1, -2, \ldots$, we choose an integer n so that $-n < \Re(\alpha) < -n + 1$, perform n integrations by parts and obtain

$$f(s) = e^{-s}\left(\frac{1}{\Gamma(\alpha+1)} + \frac{s}{\Gamma(\alpha+2)} + \cdots \frac{s^{n-1}}{\Gamma(\alpha+n)}\right)$$

$$+ \frac{s^n}{\Gamma(\alpha+n)}\int_1^\infty e^{-st}t^{\alpha+n-1}\,dt, \qquad \Re(s) > 0.$$

TABLE OF COMPARISON FUNCTIONS

α, β are complex numbers, n is a positive integer.

	$\phi(s)$	$\Phi(t)$	Restrictions
Case 1	$\frac{1}{s^\alpha}$	$\frac{t^{\alpha-1}}{\Gamma(\alpha)}$	$\alpha \neq 0, -1, -2, \ldots$
Case 2a	$\frac{\log^n s}{s^\alpha}$	$\left(-\frac{d}{d\alpha}\right)^n \frac{t^{\alpha-1}}{\Gamma(\alpha)}$	$\alpha \neq 0, -1, -2, \ldots$
Case 2b ($n = 1$)	$\frac{\log s}{s^\alpha}$	$-\frac{t^{\alpha-1}}{\Gamma(\alpha)}\left(\log t - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right)$	$\alpha \neq 0, -1, -2, \ldots$
Case 2c ($n = 1$) ($\alpha \to 0, -1, \ldots$)	$\frac{\log s}{s^\alpha}$	$(-1)^{\alpha+1}(-\alpha)!t^{\alpha-1}$	$\alpha = 0, -1, -2, \ldots$
Case 3	$\frac{e^{-\beta/s}}{s^\alpha}$	$\left(\frac{t}{\beta}\right)^{\frac{1}{2}(\alpha-1)} J_{\alpha-1}(2\sqrt{\beta t})$	none
Case 4	$\frac{e^{-\beta/s}\log^n s}{s^\alpha}$	$\left(-\frac{d}{d\alpha}\right)^n\left(\frac{t}{\beta}\right)^{\frac{1}{2}(\alpha-1)} J_{\alpha-1}(2\sqrt{\beta t})$	none
Case 5	$\frac{1}{s^\alpha \log s} - \frac{1}{s-1}$	$\int_\alpha^{\alpha+\infty} \frac{t^{u-1}}{\Gamma(u)}\,du - e^t$	none

The first term is an entire function, and the integral in the second term can be treated as above in (4). We then obtain

$$f(s) = s^{-\alpha} + \text{an entire function}, \qquad \Re(s) > 0. \tag{5}$$

We now change the path of integration in the t-plane from the ray $(1, \infty)$ to the contour consisting of the line segment $(1, 2)$, the circular arc $(2, 2e^{i\theta})$ and the ray $(2e^{i\theta}, \infty e^{i\theta})$. We obtain in this way the analytic continuation of $f(s)$ into the half-plane $\Re(se^{i\theta}) > 0$. (To obtain the continuation into the lower half-plane we let θ increase to $\frac{1}{2}\pi$ and for the upper half-plane we let θ decrease to $-\frac{1}{2}\pi$.) Equation (3) becomes

$$\begin{aligned} f(s) &= \int_1^2 e^{-st}\frac{t^{\alpha-1}}{\Gamma(\alpha)}\,dt + 2^\alpha \int_0^\theta e^{-2se^{i\phi}}\frac{e^{i\alpha\phi}}{\Gamma(\alpha)}\,d\phi + e^{\alpha i\theta}\int_2^\infty e^{-sue^{i\theta}}\frac{u^{\alpha-1}}{\Gamma(\alpha)}\,du \\ &= s^{-\alpha} + \text{an entire function}, \qquad \Re(se^{i\theta}) > 0, \end{aligned} \tag{6}$$

by an argument similar to that used in (4) and (5). Thus $f(s)$ has the representation required by the theorem in case 1.

That $f(s)$ satisfies condition (i) of Theorem 1 follows from the representation which has been established.

Condition (iii) (and the fact that the integral defined in (iii) is the original $F(t)$) follows from the absolute convergence of (3) and the fact that $F(t)$ is of bounded variation in every finite interval and continuous except at $t = 1$ (**4**, p. 212).

For both conditions (ii) and (iv) we have first to show that $f(s) \to 0$ as $s \to \pm i\infty$. If we put $s = -iy (y > 0)$ then in (6) we can choose θ, independently of y and where $0 < \theta < \pi$, so that the middle term is less in absolute value than any $\epsilon > 0$. The third term then approaches zero as $y \to \infty$ because $\Re(se^{i\theta}) > 0$, and so does the first term because of the Riemann-Lebesgue theorem. Putting $s = iy(y > 0)$ leads to a similar result.

The remainder of condition (iv) follows from the absolute convergence of (3) and the Riemann-Lebesgue theorem.

For the remainder of condition (ii) we first choose θ so that all rays in the sector $\frac{1}{2}\pi < \arg s < \psi < \pi$ enter the half-plane $\Re(se^{i\theta}) > 0$. The third term in (6) then approaches zero as $|s| \to \infty$ in this sector; the first and second terms are easily shown to be $O[\exp(-2\Re(s))]$ as $|s| \to \infty$ in the sector. Similar results hold in the sector $-\pi < \psi < \arg s < -\frac{1}{2}\pi$ and (ii) follows for any $k > 2$.

Thus Theorem 2 has been proven for case 1 of the table. From our point of view the importance of this result is that it enables us to relate a function $\Phi(t)$ to a transform which has a singularity at the origin like $s^{-\alpha}$ *even* when $\Re(\alpha) < 0$. The essential property of $\Phi(t)$ which made the proof possible was that it satisfied the equation

$$\frac{d}{dt} F(t, \alpha) = F(t, \alpha - 1)$$

so that successive integrations by parts could be carried out to give (5). Many functions are known to satisfy this "F" equation (**7**). For example, all the functions $\Phi(t)$ which appear in the table satisfy this equation and the proof of the theorem in each case goes through as in case 1.

The functions $\Phi(t)$ in the table have been chosen because their transform functions $\phi(s)$ have essentially distinct types of singularities at the origin. In the next section we shall show how even these few choices enable us to handle a considerable variety of Laplace transforms.

Cases 2 and 4 can be obtained formally from cases 1 and 3 respectively by differentiation with respect to the parameter α. Case 5 can be obtained formally from case 1 by first integrating with respect to α and then subtracting the effect of the pole at $s = 1$. (In this case we restrict our attention to the branch on which $\log s$ is real when s is real.)

From a somewhat different point of view cases 1 and 2 have been considered by Doetsch (**4**, pp. 494–502).

4. Procedure and examples. We shall now show how the results of the preceding two sections can be used to obtain the asymptotic behaviour of the inverse of a variety of Laplace transforms. We shall consider a few examples

to illustrate the procedure and then summarize the steps to be taken in any particular case.

We shall first consider the function

$$f(s) = \frac{\log s}{1+s^2}.$$

This function is a Laplace transform which can be inverted by means of the inversion integral **(4**, pp. 263–264**)** and the problem is to obtain the behaviour of its inverse $F(t)$ for large t. (In fact $F(t) = \cos t \operatorname{Si}(t) - \sin t \operatorname{Ci}(t)$ but this need not be known in what follows.)

If we subtract from $f(s)$ the effect of the poles at $s = \pm i$ we obtain the function

$$\frac{\log s}{1+s^2} - \frac{\pi s}{2+2s^2}$$

which satisfies the conditions of Theorem 1. Therefore

$$\text{(7)} \quad F(t) - \tfrac{1}{2}\pi \cos t = \frac{1}{2\pi i}\int_C e^{st}\left[\frac{\log s}{1+s^2} - \frac{\pi s}{2+2s^2}\right] ds$$

$$= \frac{1}{2\pi i}\int_C e^{st}\left[\sum_{j=0}^{n}(-1)^j s^{2j}\log s + \frac{(-1)^{n+1}s^{2n+2}\log s}{1+s^2} - \frac{\pi s}{2+2s^2}\right] ds$$

where C in Fig. 1 may be taken to the left of the points $s = \pm i$.

Now from Theorem 2 (case 2*c*) we know that

$$\text{(8)} \quad -\sum_{j=0}^{n}(-1)^j(2j)!t^{-2j-1} = \frac{1}{2\pi i}\int_C e^{st}\left[\sum_{j=0}^{n}(-1)^j s^{2j}\log s + \text{entire function}\right] ds.$$

In each of (7) and (8) contributions to the integrals from C', the part of C outside the circle $|s| = r$, are $o(e^{-\epsilon t})$ as $t \to \infty$, for some $\epsilon > 0$ because of Theorem 1. Contributions from the remainder of C due to the terms $\pi s/(1+s^2)$ and the entire function are also $o(e^{-\epsilon t})$ as $t \to \infty$ because the contour can be moved, as far as these terms are concerned, to a position lying entirely to the left of the imaginary axis.

We therefore obtain

$$F(t) - \tfrac{1}{2}\pi\cos t + \sum_{j=0}^{n}(-1)^j(2j)!t^{-2j-1} = \frac{(-1)^{n+1}}{2\pi i}\int_{C,\,|s|<r}\frac{e^{st}s^{2n+2}\log s}{1+s^2}\,ds + o(e^{-\epsilon t})$$

as $t \to \infty$. The integral on the right can be divided into two parts: one taken along the rays $\arg s = \pm\psi$ and the other taken around a portion of a small circle enclosing the origin. We choose t large and let the radius of this circle be $1/t$. From the first part we then have

$$\left|\int e^{st}\frac{s^{2n+2}\log s}{1+s^2}\,ds\right| < A\int_{1/t}^{r} e^{ut\cos\psi}u^{2n+1}\,du$$

$$= At^{-2n-2}\int_1^{rt} e^{v\cos\psi}v^{2n+1}\,dv$$

$$< At^{-2n-2}\int_1^{\infty} e^{v\cos\psi}v^{2n+1}\,dv$$

$$= Bt^{-2n-2}$$

where A, B are finite constants. For the second part we have

$$\left| \int e^{st} \frac{s^{2n+2} \log s}{1+s^2} ds \right| < e^{t(1/t)} (1/t)^{2n+1} 2\pi (1/t).$$

We therefore obtain finally

$$F(t) = \tfrac{1}{2}\pi \cos t - \sum_{j=0}^{n} (-1)^j (2j)! t^{-2j-1} + O(t^{-2n-2})$$

as $t \to \infty$.

As a second example consider

$$\begin{aligned} f(s) &= \frac{e^{-1/s}}{s(1+s)} \\ &= \sum_{j=0}^{n} (-1)^j s^{j-1} e^{-1/s} + \frac{(-1)^{n+1} s^n e^{-1/s}}{1+s}. \end{aligned}$$

The procedure for this example is similar to the one given above except in the treatment of the remainder. This time we let the radius of the small circle about the origin be $t^{-\frac{1}{2}}$ so that, for the part of the remainder taken along the rays $\arg s = \pm \psi$, we have

$$\begin{aligned} \left| \int \frac{e^{st} e^{-1/s} s^n}{1+s} ds \right| &< C \int_{t^{-\frac{1}{2}}}^{r} e^{tu \cos \psi} e^{-\cos \psi / u} u^n du \\ &= C t^{-\frac{1}{2}n-\frac{1}{2}} \int_{1}^{r\sqrt{t}} e^{\sqrt{t} \cos \psi (v - 1/v)} v^n dv \\ &< D t^{-\frac{1}{2}n-\frac{1}{2}} \end{aligned}$$

where C, D are finite constants. For the part of the remainder taken around the small circle we obtain

$$\left| \int \frac{e^{st} e^{-1/s} s^n}{1+s} ds \right| < e^{\sqrt{t}} e^{-\sqrt{t}} t^{-\frac{1}{2}n} 2\pi t^{-\frac{1}{2}}$$

so that finally

$$F(t) = \sum_{j=0}^{n} t^{-j/2} J_j(2\sqrt{t}) + O(t^{-\frac{1}{2}n-\frac{1}{2}})$$

as $t \to \infty$.

On the other hand, for

$$f(s) = \frac{e^{1/s}}{s(1+s)}$$

we obtain

$$F(t) = \sum_{j=0}^{n} t^{-\frac{1}{2}j} I_j(2\sqrt{t}) + O(t^{-\frac{1}{2}n-\frac{1}{2}} e^{+2\sqrt{t}})$$

as $t \to \infty$.

As a final example let us consider a transform with $\log s$ in the denominator. If

$$f(s) = \frac{1}{(1-s)\log s}$$

we obtain, as we did in (7),

$$F(t) + (\tfrac{1}{2} + t)e^t$$
$$= \frac{1}{2\pi i}\int_C e^{st}\left[\sum_{j=0}^{n}\left(\frac{s^j}{\log s} - \frac{1}{s-1}\right) + \frac{n+1}{s-1} + \frac{s^{n+1}}{(1-s)\log s} + \frac{1}{2}\,\frac{s+1}{(s-1)^2}\right]ds.$$

Using case 5 of the table and treating the remainder as we did in the first example we obtain finally

$$F(t) = -(\tfrac{1}{2} + t)e^t - \sum_{j=0}^{n}\left(\int_{-j}^{\infty}\frac{t^{u-1}}{\Gamma(u)}\,du - e^t\right) + O(t^{-n-2})$$

as $t \to \infty$.

It is clear that the above procedure can be applied even when the singularity being considered is not at the origin; a simple change of variable will move the singularity to the origin and multiply the function $F(t)$ by an exponential factor.

In summary, the procedure for finding the asymptotic behaviour of the inverse of a given $f(s)$ can be described briefly as follows. We first verify that the inverse exists and can be found by means of the inversion integral (**4**, pp. 263–264). We then move the singularity under consideration to the origin, subtract the effect of any poles to the right of, or on, the imaginary axis and verify that the conditions of Theorem 1 are satisfied. We then subtract the effect of a finite number of appropriate terms chosen from the table and finally examine the size of the remainder for large t.

The procedure can also be used when we have to consider a finite number of singularities with the same real part; each singularity can be considered separately.

5. Asymptotic behaviour of a certain function. In the first example of the previous section the asymptotic expansion proceeded in negative powers of t. A comparable expansion, which will of course contain sine and cosine terms, can be obtained for the examples involving Bessel functions; we have only to substitute the known asymptotic expansions of the Bessel functions and neglect any term whose order is the same as or smaller than the order of the remainder.

To obtain an expansion in terms of elementary functions for the last example we would need an asymptotic expansion for the function

$$\Phi(t) = \int_{\alpha}^{\alpha+\infty}\frac{t^{u-1}}{\Gamma(u)}\,du - e^t.$$

We shall find an expansion of this function but it will turn out that the expansion proceeds in negative powers of $\log t$. It can therefore be substituted into examples like the last one only when the first term in the sum is sufficient.

Because of Theorems 1 and 2 we can write

$$\Phi(t) = \frac{1}{2\pi i}\int_C \frac{e^{st}}{s^{\alpha}\log s}\,ds + o(e^{-\epsilon t}), \qquad t \to \infty.$$

Letting the radius of the circular part of C be $1/t$ and $s = u/t$ we obtain

$$\Phi(t) = \frac{t^{\alpha-1}}{2\pi i}\int \frac{e^u du}{u^\alpha(\log u - \log t)} + o(e^{-\epsilon t}), \qquad t \to \infty,$$

where the integration is taken around a contour like C except that the circle about the origin has radius $=1$.

$$\Phi(t) = -\frac{t^{\alpha-1}}{\log t}\left[\frac{1}{2\pi i}\int \frac{e^u}{u^\alpha}\sum_{l=0}^{n}\left(\frac{\log u}{\log t}\right)^l du + \frac{1}{2\pi i}\int \frac{e^u}{u^\alpha}\left(\frac{\log u}{\log t}\right)^{n+1}\frac{du}{1 - \log u/\log t}\right]$$

$$= -\frac{t^{\alpha-1}}{\log t}\left[\sum_{l=0}^{n}\left(\frac{1}{\log t}\right)^l\left(-\frac{d}{d\alpha}\right)^l\frac{1}{\Gamma(\alpha)} + o\left(\left(\frac{1}{\log t}\right)^{n+1}\right)\right]$$

as $t \to \infty$. We have used Hankel's contour integral representation of $1/\Gamma(\alpha)$. Our result includes some special cases considered by Colombo (3).

In case $\Re(\alpha) > -1$

$$\Phi(t) = \frac{1}{2\pi i}\int_C \frac{e^{st}}{s^\alpha \log s}\,ds$$

for all $t > 0$. By taking $\psi = \pi$ and shrinking the circular part of C to zero we can obtain a result found by Ramanujan for real α (5, p. 196).

We wish to thank F. M. C. Goodspeed and R. D. James for helpful discussions during the preparation of this paper.

References

1. V. Bush, *Operational circuit analysis* (New York, 1929).
2. J. R. Carson, *Electric circuit theory and the operational calculus* (New York, 1926).
3. S. Colombo, *Sur quelques transcendantes introduites par la résolution des équations intégrales de Volterra à noyaux logarithmiques*, Bull. Sci. Math. (2), 77 (1953), 89–104.
4. G. Doetsch, *Handbuch der Laplace-Transformation* (Basel, 1950), Bd. 1.
5. G. H. Hardy, *Ramanujan* (Cambridge, 1940).
6. O. Heaviside, *Electromagnetic theory* (London, 1899), vol. 2.
7. C. Truesdell, *A unified theory of special functions* (Princeton, Ann. Math. Studies, no. 18, 1948).

University of British Columbia

ON THE POTENTIAL THEORY OF COCLOSED HARMONIC FORMS

G. F. D. DUFF

1. Introduction. The potential theory of real harmonic tensors, which was first studied by Hodge **(5)**, offers a variety of problems by no means all of which have yet been examined. In the present paper there are formulated the solutions of some boundary value problems for the Poisson equations associated with coclosed harmonic forms. These problems include as special cases a number of previous results on coclosed harmonic forms and harmonic fields. In turn, however, they are themselves special cases of the mixed boundary value problem for harmonic forms which was studied in a preceding paper **(3)**. The method used with these boundary value theorems for coclosed harmonic forms is also applied to the differential equations of harmonic fields and of a new special class of harmonic forms which will be called biharmonic fields.

The notations and results of the theory of harmonic p-tensors as developed in **(4)** or **(8)** will be assumed known to the reader. The theory of the mixed boundary value problem for $\Delta\phi = 0$ **(3)** will also be used. We consider throughout a finite positive definite Rumannian manifold M of class C^∞, and dimension N having a smooth boundary B of dimension $N - 1$. The data of the various problems will also be supposed sufficiently differentiable.

2. The mixed problem for Poisson's equation. Since we shall need to apply the mixed problem to non-homogeneous equations, we formulate here the necessary and sufficient condition for the existence of a solution of the mixed problems in such cases. We recall that $\Delta = d\delta + \delta d$.

LEMMA 1. *There exists a solution ϕ of $\Delta\phi = \rho$, in M, with given values of $t\phi$ and $t\delta\phi$ on B, if and only if*

$$(2.1) \qquad (\rho, \tau)_M - \int_B \delta\phi \wedge *\tau = 0$$

for every eigenform τ of the eigenspace $K = \{\tau \mid d\tau = 0, \delta\tau = 0, t\tau = 0\}$.

Proof. The necessary and sufficient condition **(9)**

$$(2.2) \qquad (\rho, \tau) = 0, \qquad d\tau = 0, \delta\tau = 0, t\tau = 0, n\tau = 0,$$

for the solvability of $\Delta\phi = \rho$ without any boundary conditions is satisfied in view of (2.1). Therefore a form ϕ_1 with Laplacian $\Delta\phi_1 = \rho$ in M exists. We now seek a harmonic form ψ which satisfies the two boundary conditions

$$t\psi = t\phi - t\phi_1, \; t\delta\psi = t\delta\phi - t\delta\phi_1.$$

Received April 30, 1954.

By the theorem **(3)**, such a form ψ exists if and only if for all $\tau \in K$ we have

$$\int_B t\delta(\phi - \phi_1) \wedge *\tau = 0. \tag{2.3}$$

But since $t\tau = 0$, we see from Green's formula that

$$\int_B t\delta\phi_1 \wedge *\tau = \int_B \delta\phi_1 \wedge *\tau - \tau \wedge *d\phi,$$
$$= -D(\phi_1, \tau) + (\tau, \Delta\phi_1) = (\tau, \rho),$$

and (2.3) is therefore equivalent to (2.1). This proves the lemma.

By adding if necessary elements of the eigenspace K to the solution ϕ of the mixed boundary value problem we can arrange that $(\phi, \tau) = 0$, $\tau \in K$. The solution is then unique. For convenience this will be done in the following work. We also recall that the dimension of K is the relative Betti number $R_p(M, B)$ of M modulo its boundary B, and that this is equal to the absolute Betti number $R_q(M)$ of the complementary dimension $q = N - p$.

Let $g_K(x, y)$ be the Green's form of degree p for the mixed problem, as defined in **(3)**. Then the solution of the above Poisson equation is given by

$$(g_K, \rho) - \int_B t\phi \wedge *dg_K - \int_B t\delta\phi \wedge *g_K. \tag{2.4}$$

In the boundary value problem dual to the above, we assign values of $n\phi$ and $nd\phi$. The condition of solvability for $\Delta\phi = \rho$ is in this case

$$(\rho, \tau)_M + \int_B \tau \wedge *d\phi = 0 \tag{2.5}$$

for all τ satisfying $d\tau = 0$, $\delta\tau = 0$, $n\tau = 0$. These eigenforms τ span the eigenspace M of dimension $R_p(M)$.

3. A Dirichlet problem for coclosed harmonic forms. The differential equations satisfied by coclosed harmonic forms are $\delta d\phi = 0$ and $\delta\phi = 0$. It was shown in **(3)** and **(4)** that there exist solutions of these equations with assigned values of $t\phi$. We now study the slightly more general problem with non-homogeneous differential equations.

THEOREM I. *Let $\rho = \rho_p$ and $\sigma = \sigma_{p-1}$ be given coderived forms defined in M and let $\theta = \theta_p$ be a p-form defined on B. Then there exists a unique solution ϕ of the equations*

$$\delta d\phi = \rho, \quad \delta\phi = \sigma, \tag{3.1}$$

satisfying the boundary condition

$$t\phi = \theta \tag{3.2}$$

and the orthogonality condition $(\phi, \tau) = 0$, $\tau \in K$.

The uniqueness follows immediately since the difference $\phi = \phi_1 - \phi_2$ of of two solutions satisfies $\Delta\phi = 0$, $t\phi = 0$, $t\delta\phi = 0$ and $(\phi, \tau) = 0$, $\tau \in K$;

and so is zero from the theory of the mixed boundary value problem for harmonic forms.

Proof. We now formulate the appropriate mixed boundary value problem

$$\Delta\phi = \rho + d\sigma, \quad t\phi = \theta, \quad t\delta\phi = t\sigma, \tag{3.3}$$

with $(\phi, \tau) = 0$ for $\tau \in K$. We first show that a solution of this problem exists. The orthogonality condition of Lemma 1 is, for $\tau \in K$, the vanishing of

$$(\rho + d\sigma, \tau) - \int_B \sigma \wedge *\tau$$
$$= (\rho, \tau)_M + (\sigma, \delta\tau)_M + \int_B \sigma \wedge *\tau - \int_B \sigma \wedge *\tau = (\rho, \tau)_M.$$

Since ρ is coderived, $\rho = \delta\pi$ say, we find

$$(\rho, \tau) = (\delta\pi, \tau) = (\pi, d\tau) - \int_B \tau \wedge *\pi = 0,$$

since $d\tau = 0$, $t\tau = 0$. Thus the condition holds and a unique solution ϕ exists.

Now let $\psi = \psi_{p-1}$ be defined as

$$\psi = \delta\phi - \sigma. \tag{3.4}$$

Then from (3.3),

$$t\psi = t\delta\phi - t\sigma = 0,$$

while, since σ is coderived, $\sigma = \delta\xi$ say, we have $\psi = \delta(\phi - \xi)$ and hence $\delta\psi = 0$. Also,

$$d\psi = d\delta\phi - d\sigma = \rho - \delta d\phi, \tag{3.5}$$

from the differential equation in (3.3). Thus

$$\delta d\psi = \delta\rho = 0$$

so that

$$N(d\psi) = (\psi, \delta d\psi)_M + \int_B \psi \wedge *d\psi = 0,$$

since $t\psi$ is zero. Hence $d\psi \equiv 0$ in M. This shows that the first of (3.1) is satisfied by ϕ. Finally,

$$N(\psi) = (\psi, \delta(\phi - \xi))_M = (d\psi, \phi - \xi) - \int_B \psi \wedge *(\phi - \xi) = 0$$

since $d\psi = 0$ and $t\psi = 0$. Therefore $\psi \equiv 0$ in M and (3.4) shows that the second equation of (3.1) holds also. This concludes the proof of Theorem I.

We next consider the conditions which must be satisfied by the data of the problem if ϕ is to have further special properties. Obviously ϕ is coclosed if σ is zero. We may however ask: when is ϕ coderived, closed, or derived? To answer the first of these questions we require the following:

LEMMA 2. *If a coclosed form α on M is orthogonal to the eigenspace K, then α is coderived.*

The proof is based on the bilinear formula for closed forms on a closed manifold (**6**, p. 85). Let F be the double of our finite manifold M (**4**). Then, if $\alpha = \alpha_p$ is coclosed and $\beta = \beta_p$ is closed on F, we may write the bilinear formula

$$(\alpha, \beta)_F = \sum_{i,j=1}^{R_p(F)} e^p_{ij}\, \omega^i \nu^j. \tag{3.6}$$

Here the matrix e^p is the transposed inverse of the intersection matrix α^p of the p-cycles of a fundamental base of F with the q-cycles of a complementary fundamental base; the ω^i are the periods of $*\alpha$ on the q-cycles; the ν^i are the periods of β on the p-cycles; and

$$q = N - p.$$

We may suppose that the cycles of the two fundamental bases have been chosen as follows. The q-cycles consist of $R_q(M)$ independent absolute q-cycles of M, forming a fundamental base for M, together with certain additional cycles of $F - M$. The p-cycles, dual to the above q-cycles of F, which lie all or partly in M constitute a base for the relative p-cycles of M, mod B. To see this, we note that any relative p-cycle of M mod B, together with its image in $F - M$, constitutes an absolute p-cycle of F which is expressible as a sum of the fundamental p-cycles of F. The intersection submatrix of these q-cycles and relative p-cycles of M is again non-singular.

Since $\tau_p \in K$ is closed, and $t\tau_p = 0$, τ has zero period over any cycle lying in a sufficiently small neighbourhood $B \times I$ of B in M. Thus τ is derived in $B \times I$ (**1**); let $\tau = d\gamma_{p-1}$ there.

Let ρ_ϵ be a C^∞ scalar function of a real variable t which satisfies

$$\rho_\epsilon = 0 \text{ for } t < 0, \quad 0 < \rho_\epsilon < 1, \quad \rho_\epsilon = 1, \text{ for } t > \epsilon.$$

Such a function is easily constructed. Now let x^N be a variable which parametrizes the interval I, uniformly over the neighbourhood $B \times I$, and define

$$\beta_\epsilon = \begin{cases} 0 & \text{in } F - M, \\ d(\rho_\epsilon(x^N)\, \gamma_{p-1}) & \text{in } B \times I, \\ \tau & \text{in } M - B \times I. \end{cases}$$

Then β_ϵ is closed, of class C^∞ in F, and vanishes outside of M. Moreover $\lim_{\epsilon \to 0} \beta_\epsilon = \tau$ in M, while for any cycle $A_p{}^i$ of F,

$$\gamma^i = \lim_{\epsilon \to 0} \nu^i(\epsilon) = \lim_{\epsilon \to 0} \int_{A_p{}^i} \beta_\epsilon = \int_{R_p{}^i} \tau,$$

where $R_p{}^i$ is the relative p-cycle of M mod B which is the part of $A_p{}^i$ lying in M.

From (3.6) we find

$$\begin{aligned}(\alpha, \tau)_M = \lim_{\epsilon \to 0} (\alpha, \beta_\epsilon)_F &= \lim_{\epsilon \to 0} \sum_{i,j}^{R_p(F)} e^p_{ij} \omega^i \nu^j(\epsilon) \\ &= \sum_{i,j}^{R_p(M,B)} e^p_{ij} \omega^i \nu^j(0).\end{aligned}$$

Thus

$$(\alpha, \tau)_M = \sum_{i,j}^{R_p(M,B)} \epsilon^p_{ij}\omega^i\nu^j,$$

where the ν^j are the relative periods of τ on a fundamental base for M mod B. But we can find a form $\tau \in K$ having assigned periods on these cycles. Since $(\alpha, \tau) = 0$, $\tau \in K$, we see that

$$\sum_{i,j}^{R_p(M,B)} \epsilon^p_{ij}\omega^i\nu^j = 0,$$

for arbitrary ν^j. Because ϵ^p_{ij} $(i, j = 1, \ldots, R_p(M, B))$ is non-singular, we conclude that $\omega^i = 0$. Thus $*\alpha$ has zero periods on the absolute q-cycles of M and so is a derived form **(1)**. We therefore conclude that α is coderived, as stated in the lemma.

Taking $\sigma = 0$ in Theorem I and applying the lemma, we have

COROLLARY Ia. *When $\sigma = 0$ the solution ϕ is coderived.*

In order that ϕ be closed it is clearly necessary to have $\rho = 0$. Also $\theta = t\phi$ must be admissible in the sense of Tucker, that is,

$$d_B\theta = 0 \text{ and } \int_{\delta R^i_{p+1}} \theta = 0,$$

for every relative $(p + 1)$-cycle R^i_{p+1}. Conversely, these conditions are sufficient, because, if they are satisfied, we see that $t\phi = \theta = t\xi$, say, where ξ is a form closed in M **(1)**. Then, since $\delta d\phi = 0$, we have

$$N(d\phi) = (\phi, \delta d\phi) + \int_B \phi \wedge *d\phi$$
$$= \int_B \xi \wedge *d\phi = (d\xi, d\phi) - (\xi, \delta d\phi) = 0,$$

and so $d\phi = 0$ in M.

COROLLARY Ib. *The form ϕ is closed if and only if ρ vanishes and θ is admissible.*

For ϕ to be a derived form it must in addition have vanishing periods on all p-cycles of M. In particular θ must be a derived p-form in B, $\theta = d\zeta$ say. The remaining conditions may then be expressed

$$\int_{R_p{}^i} \phi = \int_{bR_p{}^i} \zeta,$$

but in this formula ϕ itself still appears. Replacing ϕ by its value as given by (2.4) and (3.3) we could find the explicit conditions on the data of the problem, but these are too cumbrous to be useful.

4. A mixed problem for coclosed harmonic forms. The two mixed boundary value problems of **(3)** are equivalent under dualization. However, since the equations $\delta d\phi = 0$, $\delta\phi = 0$ of coclosed harmonic forms are not self-dual, the

two lead to different problems for coclosed harmonic forms. The second of these is

THEOREM II. *Let $\rho = \rho_p$ and $\sigma = \sigma_{p-1}$ be two coderived forms in M. Let $\xi = \xi_q$ and $\eta = \eta_{q-1}$ be forms defined on B such that*

$$d_B\xi = (-1)^N t{*}\sigma, \quad \int_{bR_{q+1}} \xi = (-1)^N \int_{R_{q+1}} {*}\sigma, \tag{4.1}$$

and

$$d_B\eta = (-1)^N t{*}\rho, \quad \int_{bR_q} \eta = (-1)^N \int_{R_q} {*}\rho. \tag{4.2}$$

Then there exists a unique p-form ϕ satisfying the differential equations

$$\delta d\phi = \rho, \qquad \delta\phi = \sigma, \tag{4.3}$$

the boundary conditions

$$t{*}\phi = \xi, \quad t{*}d\phi = \eta, \tag{4.4}$$

and the orthogonality condition $(\phi, \tau) = 0$, $\tau \in M$.

The conditions (4.1) and (4.2) are necessary consequences of (4.3) and (4.4); we wish to show their sufficiency.

Proof. Consider the M-problem (3)

$$\Delta\phi = \rho + d\sigma, \quad t{*}\phi = \xi, \quad t{*}d\phi = \eta; \quad (\phi, \tau) = 0, \ \tau \in M. \tag{4.5}$$

According to (2.5), a solution exists provided that for $\tau \in M$, the quantity

$$\begin{aligned}&(\rho + d\sigma, \tau) + \int_B \tau \wedge {*}d\phi \\ &= (\rho, \tau) + (\sigma, \delta\tau) + \int_B \sigma \wedge {*}\tau + \int_B \tau \wedge \eta = (\rho, \tau)_M + \int_B \tau \wedge \eta\end{aligned}$$

vanishes. Since ρ is coderived, $\rho = \delta\pi$ say, the condition becomes

$$0 = (\delta\pi, \tau) + \int_B \tau \wedge \eta = (\pi, d\tau) + \int_B \tau \wedge (\eta - {*}\pi).$$

The volume integral on the right vanishes since $d\tau = 0$. We shall now show that $\eta - t{*}\pi$ is an admissible tangential boundary value on B. Indeed,

$$d_B(\eta - t{*}\pi) = d_B\eta - td{*}\pi = (-1)^N t{*}\rho - (-1)^N t{*}\delta\pi = 0,$$

since $\rho = \delta\pi$. Likewise,

$$\begin{aligned}\int_{bR_q'} (\eta - {*}\pi) &= \int_{bR_q'} \eta - \int_{R_q'} d{*}\pi = \int_{bR_q'} \eta - (-1)^N \int_{R_q'} {*}\delta\pi \\ &= \int_{bR_q'} \eta - (-1)^N \int_{R_q'} {*}\rho = 0,\end{aligned}$$

from (4.2). Thus $\eta - t{*}\pi = t\alpha$ where $d\alpha = 0$ in M. In consequence we find

$$\int_B \tau \wedge (\eta - {*}\pi) = \int_B \tau \wedge \alpha = \int_M d(\tau \wedge \alpha) = 0,$$

since $d\alpha = 0$ and $d\tau = 0$. Therefore the necessary condition is satisfied and a solution ϕ of (4.5) exists.

Next we define

$$\psi = \delta\phi - \sigma = \delta(\phi - \theta),$$

where θ is such that $\sigma = \delta\theta$. Then

$$d\psi = d\delta\phi - d\sigma = \rho - \delta d\rho = \delta(\pi - d\phi)$$

by the differential equation (4.5). Hence $\delta d\psi = 0$.

From (4.5), we see that

$$\begin{aligned} t{*}d\psi &= t{*}(\rho - \delta d\phi) = (-1)^N d_B\eta - t{*}\delta d\phi \\ &= (-1)^N d_B\eta - (-1)^N t d{*}d\phi \\ &= (-1)^N\{d_B\eta - d_B t{*}d\phi\} = 0. \end{aligned}$$

Hence

$$N(d\psi) = (\psi, \delta d\psi) + \int_B \psi \wedge {*}d\psi = 0,$$

and so $d\psi \equiv 0$ in M. Next we see that

$$N(\psi) = (\psi, \delta(\phi - \theta)) = (d\psi, \phi - \theta) - \int_B \psi \wedge {*}(\phi - \theta)$$

and the volume integral contains the vanishing factor $d\psi$. We show next that $t{*}(\phi - \theta)$ is admissible on B. In fact,

$$\begin{aligned} d_B t{*}(\phi - \theta) &= d_B t{*}\phi - t d{*}\theta = d_B\xi - (-1)^N t{*}\delta\theta \\ &= d_B\xi - (-1)^N t{*}\sigma = 0, \end{aligned}$$

by (4.1). Also,

$$\begin{aligned} \int_{bR_{q+1}} {*}(\phi - \theta) &= \int_{bR_{q+1}} \xi - \int_{R_{q+1}} d{*}\theta = \int_{bR_{q+1}} \xi - (-1)^N \int_{R_{q+1}} {*}\delta\theta \\ &= \int_{bR_{q+1}} \xi - (-1)^N \int_{R_{q+1}} {*}\sigma = 0, \end{aligned}$$

by the second of (4.1). Thus $t{*}(\phi - \theta)$ is admissible and so equal to $t\beta$ for some closed form β in M. That is,

$$N(\psi) = -\int_B \psi \wedge {*}(\phi - \theta) = -\int_B \psi \wedge \beta = -\int_M d(\psi \wedge \beta) = 0$$

since $d\beta = 0$ and $d\psi = 0$. Hence $\psi \equiv 0$ and the two differential equations (4.3) are satisfied. This completes the proof of the theorem.

In order that ϕ should be coderived, it is necessary that $\sigma = 0$, that ξ should be a derived form on B, and that the absolute periods of ${*}\phi$ should vanish. This last condition cannot be expressed without the Green's form g_M which is dual to the g_K of (2.4).

When ϕ is closed, ρ and η vanish. Conversely, if ρ and η are zero, we see from (4.3) and (4.4) that $\delta d\phi = 0$ and $t{*}d\phi = 0$. Thus

$$N(d\phi) = (\phi, \delta d\phi) + \int_B \phi \wedge {*}d\phi = 0$$

and ϕ is closed. Now the dual of Lemma 2 states that a closed form orthogonal to M is derived.

COROLLARY IIa. *The solution ϕ is derived if and only if ρ and η vanish.*

We also state as a corollary the special case when the differential equations are homogeneous. If ρ and σ are zero the conditions (4.1) and (4.2) show that ξ and η are admissible on B. Since ϕ is coclosed the periods of $*\phi$ are defined, and by adding a suitable eigenform $\tau \in M$ the relative period of $*\phi$ can be given assigned values on given relative q-cycles R_q^i. These periods will depend only on the bB_q^i.

COROLLARY IIb. *There exists a unique coclosed harmonic p-form ϕ such that $t*\phi$ and $t*d\phi$ have given admissible boundary values, and $*\phi$ has given periods on $R_p(M)$ independent relative q-cycles whose boundaries are fixed.*

5. The Poisson equations associated with harmonic fields. The differential equations satisfied by harmonic fields are the self-dual pair $d\phi = 0$, $\delta\phi = 0$. It has been shown that there exists a unique harmonic field having a given admissible tangential boundary value and given relative periods **(4)**. We therefore formulate the following

THEOREM III. *Let $\rho = \rho_{p+1}$ and $\sigma = \sigma_{p-1}$ be given forms, derived and coderived on M, respectively. Let ξ be a form given on B such that*

$$d_B\xi = t\rho, \quad \int_{bR^i_{p+1}} \xi = \int_{R_{p+1}} \rho \tag{5.1}$$

for all relative $(p + 1)$-cycles. Then there exists a unique form ϕ satisfying the differential equations

$$d\phi = \rho, \quad \delta\phi = \sigma, \tag{5.2}$$

the boundary condition

$$t\phi = \xi \tag{5.3}$$

and the orthogonality condition $(\phi, \tau) = 0$, $\tau \in K$.

For the proof we consider the problem

$$\Delta\phi = \delta\rho + d\sigma, \quad t\rho = \xi, \quad t\delta\phi = t\sigma; \quad (\phi, \tau) = 0, \ \tau \in K. \tag{5.4}$$

A solution exists if and only if for each $\tau \in K$,

$$(\delta\rho + d\sigma, \tau) - \int_B \sigma \wedge *\tau = 0.$$

However a simple calculation using Green's formula and the fact that $t\tau = 0$ shows that this condition is in fact satisfied. Thus a solution ϕ of (5.4) exists.

Again we define

$$\psi = \delta\phi - \sigma = \delta(\phi - \alpha), \tag{5.5}$$

since σ is derived, $\sigma = \delta\alpha$ say. We find as before

$$d\psi = d\delta\phi - d\sigma = \delta\rho - \delta d\phi$$

so that $\delta d\psi = 0$. In addition, $t\psi = t\delta\phi - t\sigma = 0$, by (5.4). Thus

$$N(d\psi) = (\psi, \delta d\psi) + \int_B \psi \wedge *d\psi = 0,$$

and we find that $d\psi \equiv 0$. Again,

$$N(\psi) = (\psi, \delta(\phi - \alpha)) = (d\psi, \phi - \alpha) - \int_B \psi \wedge *(\phi - \alpha) = 0,$$

so that ψ vanishes identically. This shows that the second of (5.2) holds.

Let $\chi = d\phi - \rho$; then $\delta\chi = \delta d\phi - \delta\rho = d\sigma - d\delta\phi = 0$, since $\delta\phi = \sigma$. Since ρ is derived, $\rho = d\zeta$ say, we find $\chi = d(\phi - \zeta)$. Now

$$N(\chi) = (\chi, d(\phi - \zeta)) = (\delta\chi, \phi - \zeta) + \int_B (\phi - \zeta) \wedge *\chi.$$

The volume integral vanishes since $\delta\chi = 0$. Now $t(\phi - \zeta)$ is an admissible boundary value since, first,

$$d_B(t\phi - t\zeta) = d_B t\phi - td\zeta = d_B\xi - t\rho = 0,$$

from (5.1), and second,

$$\int_{\delta R_{p+1}} t(\phi - \zeta) = \int_{\delta R_{p+1}} t\phi - \int_{R_{p+1}} d\zeta = \int_{\delta R_{p+1}} \xi - \int_{R_{p+1}} \rho = 0,$$

by (5.4) and (5.1). Hence $t(\phi - \zeta) = t\gamma$ say, where $d\gamma = 0$ in M, and

$$\int_B (\phi - \zeta) \wedge *\chi = \int_B \gamma \wedge *\chi = \int_M d(\gamma \wedge *\chi) = 0,$$

since $d\gamma = 0$, $\delta\chi = 0$. Thus $\chi \equiv 0$ in M and so the equations (5.2) are both valid. This completes the proof of the theorem.

From Lemma 2 it is evident that the solution normalized orthogonal to K is coderived if $\sigma = 0$. This problem has been studied in Euclidean space by Miranda **(7)**.

6. A mixed problem for biharmonic fields. We consider here a new class of harmonic forms, which satisfy the self-dual equations $\delta d\phi = 0$, $d\delta\phi = 0$. Since these have the same relation to biharmonic forms ($\Delta^2\phi = 0$) as do harmonic fields to harmonic forms, they shall be called biharmonic fields. The mixed problem for harmonic forms leads to the following result on biharmonic fields.

THEOREM IV. *Let ρ and σ be forms of degree p on M, coderived and derived respectively. Let $\xi = \xi_p$ and $\eta = \eta_{p-1}$ be defined on B, with*

(6.1) $$d_B\eta = t\sigma, \quad \int_{\delta R_p{}^i} \eta = \int_{R_p{}^i} \sigma.$$

Then there exists a unique form $\phi = \phi_p$ satisfying the differential equations

(6.2) $$\delta d\phi = \rho, \quad d\delta\phi = \sigma,$$

the boundary conditions

(6.3) $$t\phi = \xi, \quad t\delta\phi = \eta,$$

and the orthogonality condition $(\phi, \tau) = 0$, $\tau \in K$.

Proof. The uniqueness can be verified at once, from Lemma 1. To establish the existence of a solution, we set up the mixed problem

$$\text{(6.4)} \qquad \Delta\phi = \rho + \sigma, \quad t\phi = \xi, \quad t\delta\phi = \eta; \quad (\phi, \tau) = 0, \ \tau \in K.$$

A solution exists if and only if for all $\tau \in K$,

$$(\rho + \sigma, \tau) - \int_B \eta \wedge *\tau$$

vanishes. Since ρ is coderived, $\rho = \delta\theta$ say, and also $\sigma = d\zeta$; this quantity is equal to

$$\begin{aligned} &(\delta\theta + d\zeta, \tau) - \int_B \eta \wedge *\tau \\ &= (\theta, d\tau) - \int_B \tau \wedge *\theta + (\zeta, \delta\tau) + \int_B \zeta \wedge *\tau - \int_B \eta \wedge *\tau \\ &= \int_B (\zeta - \eta) \wedge *\tau. \end{aligned}$$

We show that $t\zeta - \eta$ is admissible. Indeed,

$$d_B(t\zeta - \eta) = td\zeta - d_B\eta = t\sigma - d_B\eta = 0,$$

by the first of (6.1). Then also

$$\int_{bB_p{}^i} (\zeta - \eta) = \int_{B_p{}^i} d\zeta - \int_{bB_p{}^i} \eta = \int_{B_p{}^i} \sigma - \int_{bB_p{}^i} \eta = 0,$$

by the second of (6.1). Hence $t\zeta - \eta = t\alpha$, where $d\alpha = 0$ in M, and

$$\int_B (\zeta - \eta) \wedge *\tau = \int_B \alpha \wedge *\tau = \int_M d(\alpha \wedge *\tau) = 0,$$

as in previous calculations. Thus the condition relative to (6.4) is satisfied, and a solution ϕ exists. It remains to be shown that (6.2) are satisfied.

Now let

$$\text{(6.5)} \qquad \psi = \delta d\phi - \rho = \sigma - d\delta\phi.$$

Since $\rho = \delta\theta$, we find that $\psi = \delta(d\phi - \theta)$ is coderived. Also

$$t\psi = t\sigma - td\delta\phi = t\sigma - d_B t\delta\phi = t\sigma - d_B\eta = 0,$$

from (6.1). Thus we find

$$N(\psi) = (\psi, \delta(d\phi - \theta)) = (d\psi, d\phi - \theta) - \int_B \psi \wedge *(d\phi - \theta).$$

The surface integral vanishes since $t\psi = 0$. Now $d\psi = d\sigma - dd\delta\phi = 0$ since σ is derived. Thus $N(\psi) = 0$, so ψ vanishes identically. That is, the equations (6.2) are both valid. This completes the proof.

We note two special cases of Theorem IV. First, let ρ vanish so that $\delta d\phi = 0$, and let ξ be an admissible tangential boundary value. It then follows easily from Green's formula that ϕ is closed. Second, let σ and η both vanish. Green's formula then shows that ϕ is coclosed, and from Lemma 2 it follows that ϕ is coderived.

7. A Neumann problem for biharmonic fields. The Neumann boundary value problem for harmonic forms **(2)** yields an independent result in connection with biharmonic fields. We state first the Neumann theorem in the non-homogeneous case.

LEMMA 3. *There exists a unique solution ϕ of*

$$\Delta\phi = \rho \tag{7.1}$$

which satisfies the boundary conditions

$$t{*}d\phi = \xi, \quad t\delta\phi = \eta \tag{7.2}$$

and the orthogonality condition

$$(\rho, \tau) = 0, \qquad d\tau = 0, \quad \delta\tau = 0, \tag{7.3}$$

if and only if for every harmonic field τ we have

$$(\rho, \tau) - \int_B (\tau \wedge \xi - \eta \wedge *\tau) = 0. \tag{7.4}$$

The proof is similar to that of Lemma 1 and will therefore be omitted.

We state the application to the biharmonic field equations as follows.

THEOREM V. *Let ρ and σ be p-forms defined on M which are coderived and derived, respectively. Let $\xi = \xi_{q+1}$ and $\eta = \eta_{p-1}$ be forms defined on B such that*

$$d_B\xi = (-1)^N t{*}\rho, \quad \int_{bR_q{}^i} \xi = (-1)^N \int_{R_q{}^i} *\rho, \tag{7.5}$$

and

$$d_B\eta = t\sigma, \qquad \int_{bR_p{}^i} \eta = \int_{R_p{}^i} \sigma. \tag{7.6}$$

Then there exists a unique p-form ϕ satisfying the differential equations

$$\delta d\phi = \rho, \quad d\delta\phi = \sigma, \tag{7.7}$$

the boundary conditions

$$t{*}d\phi = \xi, \quad t\delta\phi = \eta, \tag{7.8}$$

and the orthogonality condition

$$(\phi, \tau)_M = 0 \tag{7.9}$$

for every harmonic field τ defined on M.

Proof. We formulate the problem

$$\Delta\phi = \rho + \sigma, \quad t{*}d\phi = \xi, \quad t\delta\phi = \eta; \quad (\phi, \tau) = 0 \text{ if } d\tau = 0, \delta\tau = 0. \tag{7.10}$$

According to Lemma 3, a solution exists if and only if for every harmonic field τ, the quantity

$$(\rho + \sigma, \tau) + \int_B (\tau \wedge \xi - \eta \wedge *\tau)$$

vanishes. Writing $\rho = \delta\theta$ and $\sigma = d\zeta$ in view of our hypotheses, and making use of Green's formula we find this expression can be put in the form

$$\int_B [\tau \wedge (\xi - *\theta) - (\eta - \zeta) \wedge *\tau]. \tag{7.11}$$

As in our previous work, we may show that the conditions (7.5) imply that $\xi - *\theta$ is admissible; and that (7.6) imply that $\eta - \zeta$ is admissible. A further application of Stokes' formula then shows that the integral (7.11) vanishes for all harmonic fields τ. Thus the condition of Lemma 3 is satisfied and a solution of (7.10) exists.

As in §6 we set $\psi = \delta d\phi - \rho = \sigma - d\delta\phi$. It again follows that $\psi = \delta(d\phi - \theta)$, and, from the first of (7.6), that $t\psi = 0$. Since $d\psi = 0$, $\delta\psi = 0$, we can again show that ψ vanishes identically. This shows that (7.7) holds and establishes Theorem V.

In contrast to Theorem IV, this last result is self-dual. When ρ and ξ vanish, we see by Green's formula that ϕ is closed, and, by the dual of Lemma 1 (since ϕ is orthogonal to the eigenspace $d\tau = 0$, $\delta\tau = 0$, $n\tau = 0$) that ϕ is derived. Dually, ϕ is coderived if σ and η are zero.

8. Concluding remarks. Four of the five theorems have been deduced from the mixed boundary value theorem for harmonic forms, and one from the Neumann theorem. Corresponding special cases of the Dirichlet theorem for harmonic forms are not known.

The conditions on the data of the theorems for harmonic forms all involve the eigenforms which are harmonic fields having some special properties. However, in the special cases considered in this paper, all of the conditions are expressible directly in terms of the data without the appearance of any class of eigenforms. The normalizations such as $(\phi, \tau) = 0$, $\tau \in K$ could be discarded, and the solutions would then lose the property of uniqueness. Or they could be replaced by suitable conditions on the periods of the solutions.

References

1. G. F. D. Duff, *Differential forms in manifolds with boundary*, Ann. Math., *56* (1952), 115–127.
2. ———, *Boundary value problems associated with the tensor Laplace equation*, Can. J. Math., *5* (1953), 196–210.
3. ———, *A tensor boundary value problem of mixed type*, Can. J. Math., *6* (1954), 427–440.
4. G. F. D. Duff and D. C. Spencer, *Harmonic tensors on Riemannian manifolds with boundary*, Ann. Math., *56* (1952), 128–156.
5. W. V. D. Hodge, *A Dirichlet problem for harmonic functionals*, Proc. London Math. Soc. (2), *36* (1933), 257–303.
6. ———, *Theory and application of harmonic integrals* (2nd ed., Cambridge, 1951).
7. C. Miranda, *Sull' integrazione delle forme differenziali esterne*, Ricerche di Math. dell' University di Napoli, II, (1953), 151–182.
8. G. de Rham and K. Kodaira, *Harmonic integrals;* mimeographed notes (Princeton, Institute for Advanced Study, 1950).
9. D. C. Spencer, *Real and complex operators on manifolds* (Contributions to the theory of Riemann surfaces, Princeton, 1953).

University of Toronto

ON SELBERG'S LEMMA FOR ALGEBRAIC FIELDS

R. G. AYOUB

1. Introduction. Recently two Japanese authors **(1)** gave a beautifully simple proof of Selberg's fundamental lemma in the theory of distribution of primes.[1] The proof is based on a curious twist in the Möbius inversion formula. The object of this note is to show that their proof may be extended to a proof of the result for algebraic fields corresponding to Selberg's lemma. Shapiro **(2)** has already derived this result using Selberg's methods and deduced as a consequence the prime ideal theorem.

Let K be an algebraic extension of the rationals of degree k, and denote by $N(\mathfrak{a})$ the norm of the ideal $\mathfrak{a}$ and by $\mathfrak{p}$, $\mathfrak{p}_i$ etc., prime ideals of K.

We define $\mu(\mathfrak{a})$ and $\Lambda(\mathfrak{a})$ as in the case of the rational field, viz.,

$$\mu(\mathfrak{a}) = \begin{cases} 1 & \text{if } \mathfrak{a} = 1 \\ (-1)^r & \text{if } \mathfrak{a} = \mathfrak{p}_1 \ldots \mathfrak{p}_r, \text{ the } \mathfrak{p}_i \text{ all different,} \\ 0 & \text{otherwise;} \end{cases}$$

$$\Lambda(\mathfrak{a}) = \begin{cases} \log N(\mathfrak{a}) & \text{if } \mathfrak{a} \text{ is a power of a prime ideal } \mathfrak{p}. \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to deduce that

$$\sum_{\mathfrak{b}|\mathfrak{a}} \mu(\mathfrak{b}) = \begin{cases} 0 & \text{if } \mathfrak{a} \neq 1 \\ 1 & \text{if } \mathfrak{a} = 1 \end{cases} \tag{1}$$

and

$$\sum_{\mathfrak{b}|\mathfrak{a}} \Lambda(\mathfrak{b}) = \log N(\mathfrak{a}). \tag{2}$$

The Möbius inversion formula is valid, i.e. if

$$f(\mathfrak{a}) = \sum_{\mathfrak{b}|\mathfrak{a}} g(\mathfrak{b})$$

then

$$g(\mathfrak{a}) = \sum_{\mathfrak{b}|\mathfrak{a}} \mu(\mathfrak{b}) f\left(\frac{\mathfrak{a}}{\mathfrak{b}}\right).$$

It follows that

$$\begin{aligned} \Lambda(\mathfrak{a}) &= \sum_{\mathfrak{b}|\mathfrak{a}} \mu(\mathfrak{b}) \log\left(\frac{N(\mathfrak{a})}{N(\mathfrak{b})}\right) \\ &= -\sum_{\mathfrak{b}|\mathfrak{a}} \mu(\mathfrak{b}) \log N(\mathfrak{b}). \end{aligned} \tag{3}$$

Define

$$\psi(x) = \sum_{N(\mathfrak{a})<x} \Lambda(\mathfrak{a}).$$

Received April 3, 1954.

[1]This proof was brought to my attention by Dr. Leo Moser of the University of Alberta.

It is our object to give a new proof of

SELBERG'S LEMMA:

$$\psi(x)\log x + \sum_{N(\mathfrak{a})<x} \Lambda(\mathfrak{a})\,\psi\left(\frac{x}{N(\mathfrak{a})}\right) = 2x\log x + O(x).$$

The proof is based on the next theorem which is the essence of the Japanese method; a factor $\log x$ is introduced in the Möbius transform with interesting consequences.

THEOREM 1.1. *If*

$$f(x) = \sum_{N(\mathfrak{a})<x} h\left(\frac{x}{N(\mathfrak{a})}\right)\log x$$

then

$$\sum_{N(\mathfrak{a})<x} \mu(\mathfrak{a})\, f\left(\frac{x}{N(\mathfrak{a})}\right) = h(x)\log x + \sum_{N(\mathfrak{a})<x} \Lambda(\mathfrak{a})\, h\left(\frac{x}{N(\mathfrak{a})}\right). \tag{4}$$

Proof.

$$\begin{aligned}\sum_{N(\mathfrak{a})<x} \mu(\mathfrak{a})\, f\left(\frac{x}{N(\mathfrak{a})}\right) &= \sum_{N(\mathfrak{a})<x} \mu(\mathfrak{a}) \sum_{N(\mathfrak{b})<x/N(\mathfrak{a})} h\left(\frac{x}{N(\mathfrak{a})\,N(\mathfrak{b})}\right)\log\frac{x}{N(\mathfrak{b})} \\ &= \sum_{N(\mathfrak{c})<x} h\left(\frac{x}{N(\mathfrak{c})}\right)\sum_{\mathfrak{b}\mid\mathfrak{c}} \mu(\mathfrak{b})\log\left(\frac{x}{N(\mathfrak{b})}\right) \\ &= h(x)\log x + \sum_{N(\mathfrak{c})<x} h\left(\frac{x}{N(\mathfrak{c})}\right)\Lambda(\mathfrak{c}),\end{aligned}$$

by (1) and (3).

2. Some estimates. We make the following abbreviation: we denote simply by the index a summation over the range 0 to x, for example,

$$\sum_{\mathfrak{a}} f(\mathfrak{a}) \text{ means } \sum_{N(\mathfrak{a})<x} f(\mathfrak{a}), \text{ while } \sum_{n} f(n) \text{ means } \sum_{n<x} f(n).$$

In other cases the range of summation will be specified. We sometimes use the notation $A \ll B$ to mean $A = O(B)$. We assume known the following classical result of Weber **(3)**:

$$[x] = \sum_{\mathfrak{a}} 1 = g\,x + a(x), \tag{5}$$

where $a(x) = O(x^{1-m})$ with $m = 1/k$, g is the residue of $\zeta_k(s)$ at $s = 1$, i.e.,

$$g = \frac{2^{r_1+r_2}\pi^{r_2}R}{w\sqrt{|d|}}\,h.$$

Here r_1 and r_2 are the numbers of real and pairs of complex conjugate fields, w is the order of the group of roots of unity, d is the discriminant, R the regulator and h is the class number.

THEOREM 2.1.

$$\sum_{\mathfrak{a}} N(\mathfrak{a})^{-1} = g\log x + c + O(x^{-m}).$$

Proof. Using (5), we get

$$\begin{aligned}
\sum_{\mathfrak{a}} N(\mathfrak{a})^{-1} &= \sum_{n} \frac{[n]-[n-1]}{n} \\
&= \sum_{n} \frac{gn-g(n-1)}{n} + \sum_{n} \frac{a(n)-a(n-1)}{n} \\
&= g\sum_{n} \frac{1}{n} + \sum_{n} a(n)\left(\frac{1}{n} - \frac{1}{n+1}\right) + O(x^{-m}) \\
&= g \log x + g\gamma + O(x^{-1}) + O\left(\sum_{n=1}^{\infty} n^{-1-m}\right) + O(x^{-m}) \\
&= g \log x + g\gamma + O(1) + O(x^{-m}) \\
&= g \log x + c + O(x^{-m}),
\end{aligned}$$

where γ is Euler's constant.

THEOREM 2.2.

$$\sum_{\mathfrak{a}} N(\mathfrak{a})^{v-1} = O(x^{v}) \qquad \text{if } 0 < v < 1.$$

Proof. Using (5) again,

$$\begin{aligned}
\sum_{\mathfrak{a}} N(\mathfrak{a})^{v-1} &= \sum_{n} ([n]-[n-1])\, n^{v-1} \\
&\ll \sum_{n} n^{v-1} + \sum_{n} \{a(n) - a(n-1)\}\, n^{v-1} \\
&\ll x^{v} + \sum_{n} n^{v-m}\left\{1 - \left(1+\frac{1}{n}\right)^{v-1}\right\} \\
&\ll x^{v} + \sum_{n} n^{v-m-1} \\
&\ll x^{v} + x^{v-m}\log x \\
&\ll x^{v}.
\end{aligned}$$

THEOREM 2.3.

$$\sum_{\mathfrak{a}} \log N(\mathfrak{a}) = g\, x \log x - g\, x + O(x^{1-m}\log x).$$

Proof. By (5),

$$\begin{aligned}
\sum_{\mathfrak{a}} \log N(\mathfrak{a}) &= \sum_{n} ([n]-[n-1]) \log n \\
&= g\sum_{n} \log n + \sum_{n} \{a(n) - a(n-1)\} \log n.
\end{aligned}$$

The second sum, however, is

$$\begin{aligned}
&\ll \sum_{n} a(n) \log\left(1+\frac{1}{n}\right) + x^{1-m}\log x \\
&\ll \sum_{n} n^{-m} + x^{1-m}\log x \\
&\ll x^{1-m}\log x.
\end{aligned}$$

Consequently,

$$\sum_{\mathfrak{a}} \log N(\mathfrak{a}) = g\, x \log x - g\, x + O(x^{1-m}\log x).$$

The object of the next paragraph is to prove

THEOREM 2.4.

$$\sum_{\mathfrak{a}} \frac{\Lambda(\mathfrak{a})}{N(\mathfrak{a})} = \log x + O(1).$$

Shapiro's proof is based on several auxiliary results which are needed for the proof of Selberg's lemma. We prove the theorem here directly, using Chebychev's ideas. We first notice that

$$\sum_{\mathfrak{a}} \frac{\Lambda(\mathfrak{a})}{N(\mathfrak{a})} = \sum_{\mathfrak{p}} \frac{\log N(\mathfrak{p})}{N(\mathfrak{p})} + \sum_{\mathfrak{p}^2} \frac{\log N(\mathfrak{p})}{N(\mathfrak{p})^2} + \dots$$
$$= \sum_{\mathfrak{p}} \frac{\log N(\mathfrak{p})}{N(\mathfrak{p})} + O(1).$$

It is therefore enough to show that the sum on the right is $\log x + O(1)$. The number of ideals $\mathfrak{a}$ with $N(\mathfrak{a}) < x$ and divisible by a prime ideal $\mathfrak{p}$ is $[x/N(\mathfrak{p})]$ and so on for $\mathfrak{p}^2$ etc. Hence

$$\prod_{\mathfrak{a}} N(\mathfrak{a}) = \prod_{\mathfrak{p}} N(\mathfrak{p})\left[\frac{x}{N(\mathfrak{p})}\right] + \left[\frac{x}{N(\mathfrak{p})^2}\right] + \dots$$

and

$$\sum_{\mathfrak{a}} \log N(\mathfrak{a}) = \sum_{\mathfrak{p}} \log N(\mathfrak{p})\left\{\left[\frac{x}{N(\mathfrak{p})}\right] + \left[\frac{x}{N(\mathfrak{p})^2}\right] + \dots\right\} \tag{6}$$
$$= \sum_{\mathfrak{p}} \log N(\mathfrak{p})\left[\frac{x}{N(\mathfrak{p})}\right]$$
$$+ O(x) \sum_{\mathfrak{p}} \log N(\mathfrak{p})\{N(\mathfrak{p})^{-2} + N(\mathfrak{p})^{-3} + \dots\}$$
$$= g\,x \sum_{\mathfrak{p}} \frac{\log N(\mathfrak{p})}{N(\mathfrak{p})} + O(x^{1-m}) \sum_{\mathfrak{p}} \frac{\log N(\mathfrak{p})}{N(\mathfrak{p})^{1-m}}$$
$$+ O(x) \sum_{\mathfrak{p}} \frac{\log N(\mathfrak{p})}{N(\mathfrak{p})^2}.$$

The third sum on the right is $O(1)$; we now evaluate the second one. For this purpose we introduce the function $\theta(x) = \sum_{\mathfrak{p}} \log N(\mathfrak{p})$. Since $N(\mathfrak{p})$ is at most p^k for some rational prime p, we conclude that $\theta(x)$ is $O(\sum_p \log p) = O(x)$. Hence

$$\sum_{\mathfrak{p}} \frac{\log N(\mathfrak{p})}{N(\mathfrak{p})^{1-m}} = \sum \frac{\theta(n) - \theta(n-1)}{n^{1-m}}$$
$$\ll \sum_{n} \theta(n)\{n^{m-1} - (n+1)^{m-1}\} + x^m$$
$$\ll \sum_{n} n^{m-1} + x^m$$
$$\ll x^m.$$

Using Theorem 2.3 and (6), we deduce Theorem 2.4.

THEOREM 2.5.

$$\sum_{\mathfrak{a}} \psi\left(\frac{x}{N(\mathfrak{a})}\right) = g\,x \log x - g\,x + O(x^{1-m}\log x).$$

Proof.

$$\sum_{\mathfrak{a}} \psi\left(\frac{x}{N(\mathfrak{a})}\right) = \sum_{\mathfrak{a}} \sum_{\mathfrak{ab}} \Lambda(\mathfrak{b})$$
$$= \sum_{\mathfrak{c}} \sum_{\mathfrak{b}|\mathfrak{c}} \Lambda(\mathfrak{b})$$
$$= \sum_{\mathfrak{c}} \log N(\mathfrak{c})$$
$$= g\,x \log x - g\,x + O(x^{1-m} \log x),$$

using (2) and Theorem 2.3.

3. Proof of Selberg's Lemma. In (4), we put $h(x) = \psi(x) - x + c/g + 1$, where c is the constant of Theorem 2.1. Then

$$h(x) \log x + \sum_{\mathfrak{a}} \Lambda(\mathfrak{a})\, h\left(\frac{x}{N(\mathfrak{a})}\right)$$
$$= \log x \left\{ \psi(x) - x + \frac{c}{g} + 1 \right\} + \sum_{\mathfrak{a}} \Lambda(\mathfrak{a})\, \psi\left(\frac{x}{N(\mathfrak{a})}\right)$$
$$- x \sum_{\mathfrak{a}} \frac{\Lambda(\mathfrak{a})}{N(\mathfrak{a})} + O(\psi(x))$$
$$= \log x\, \psi(x) + \sum_{\mathfrak{a}} \Lambda(\mathfrak{a})\, \psi\left(\frac{x}{N(\mathfrak{a})}\right) - 2x \log x + O(x) + O(\psi(x)),$$

by Theorem 2.4. On the other hand,

$$f(x) = \log x \left[\sum_{\mathfrak{a}} \psi\left(\frac{x}{N(\mathfrak{a})}\right) - x \sum_{\mathfrak{a}} N(\mathfrak{a})^{-1} + \left(\frac{c}{g} + 1\right) \sum_{\mathfrak{a}} 1 \right]$$
$$= \log x \{ g\,x \log x - g\,x + O(x^{1-m} \log x) - g\,x \log x$$
$$- cx - O(x^{-m+1}) + \left(\frac{c}{g} + 1\right)(g\,x - O(x^{1-m}))\}$$
$$= O(x^{1-m} \log^2 x) = O(x^{1-\frac{1}{2}m}),$$

by (5) and Theorems 2.1 and 2.5. Consequently

$$\sum_{\mathfrak{a}} \mu(\mathfrak{a})\, f\left(\frac{x}{N(\mathfrak{a})}\right) = O(x^{1-\frac{1}{2}m}) \sum_{\mathfrak{a}} N(\mathfrak{a})^{-1+\frac{1}{2}m}$$
$$= O(x),$$

by Theorem 2.2.

Combining these results, we conclude that

$$\psi(x) \log x + \sum_{\mathfrak{a}} \Lambda(\mathfrak{a})\, \psi\left(\frac{x}{N(\mathfrak{a})}\right) = 2x \log x + O(x) + O(\psi(x)).$$

Since $\theta(x) = O(x)$, then $\psi(x) = O(x)$, but it will be noticed that this fact is a consequence of the above inequality. The proof of the theorem is therefore complete.

REFERENCES

1. T. Tatuzawa and K. Iseki, *On Selberg's elementary proof of the prime number theorem*, Proc. Jap. Acad., 27 (1951), 340–342.
2. H. Shapiro, *An elementary proof of the prime ideal theorem*, Communications on Pure and Applied Mathematics, 2 (1949), 309–323.
3. E. Landau, *Einführung in die elementare und analytische Theorie der algebraischen Zahlen* (New York, 1949), 124 ff.

Pennsylvania State College

CORRECTION TO "ON A GENERALIZATION OF THE FIRST CURVATURE OF A CURVE IN A HYPERSURFACE OF A RIEMANNIAN SPACE"

T. K. PAN

The author takes this opportunity to correct some errors in his paper "*On a generalization of the first curvature of a curve in a hypersurface of a Riemannian Space*", in this Journal, *6* (1954), 210–216.

Page 211, line -9: for "is the unit first curvature vector of C in V_n at P" read "is a unit vector".

Page 212, line 8: for "K_g" read "$\bar{K}_g$".

9: for "K_g is the first" read "$\bar{K}_g = g_{ij} p^i \mu^j$ is called the projected first".

11: for "$K_g = 0$" read "$K_g = \sqrt{(g_{ij} p^i p^j)} = 0$".

15: for "its first" read "its projected first".

Page 215, Theorem 4.2: for "vector" and "vectors" read "unit vector" and "unit vectors".

University of Oklahoma